DOCUMENT RESUME

ED 218 036 RC 013 432

AUTHOR Nielsen, Francois; Fernandez, Roberto M.

TITLE Achievement of Hispanic Students in American High Schools: Background Characteristics and Achievement. Contractor Report.
INSTITUTION National Opinion Research Center, Chicago, Ill.
SPONS AGENCY National Center for Education Statistics (ED), Washington, DC.
REPORT NO NCES-82-219
PUB DATE Nov 81
CONTRACT OE-300-78-0208
NOTE 125p.

EDRS PRICE MF01/PC05 Plus Postage.
DESCRIPTORS *Academic Achievement; *Academic Aspiration; Comparative Analysis; Cubans; Ethnic Origins; Grade 10; High Schools; High School Seniors; *High School Students; *Hispanic Americans; Language Proficiency; *Language Usage; Latin Americans; Longitudinal Studies; Mexican Americans; Puerto Ricans; Questionnaires; Scores; Socioeconomic Status
IDENTIFIERS High School and Beyond (NCES); *Length of Residence

ABSTRACT

Presenting data and analyses from the first wave of the study "High School and Beyond" (HS&B), a longitudinal study of United States high school sophomores and seniors, the report focuses on the composition of the Hispanic tenth and twelfth grade student population with respect to various characteristics; i.e., language use and proficiency, length of U.S. residence, family's socioeconomic status. The first section describes differences among Hispanic subgroups, and between Hispanics and the rest of the student population, with respect to both measures of achievement (school delay, aspirations, test scores) and possible explanatory factors (language usage, socioeconomic status, immigration history). The next section presents a basic model of the educational achievement process for Hispanics to investigate the relative importance of the basic input factor (language use and proficiency, family socioeconomic status, length of U.S. residence) as determinants of Hispanic educational attainment, and to explore whether differences in achievement profiles of Hispanic subgroups, which are previously noted, can be explained by subgroup differences in these basic input factors. Prior to presentation of the results, the HS&B data, selection of samples, and other methodological aspects of the study are discussed. Selected questions on language use, language proficiency, and national origin are appended. (NQA)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED218036

SCOPE OF INTEREST NOTICE
The ERIC Facility has assigned this document for processing to: RC | UD
In our judgement, this document is also of interest to the clearinghouses noted to the right. Indexing should reflect their special points of view.

APR 1982
ERIC/CRESS
RECEIVED

Contractor Report

Achievement of Hispanic Students in American High Schools: Background Characteristics and Achievement

National Center for Education Statistics

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

RC 01 3492

Achievement of Hispanic Students in American High Schools: Background Characteristics and Achievement

National Opinion Research Center

Francois Nielsen
Roberto M. Fernandez

Samuel S. Peng
Project Officer
National Center for Education Statistics

November 1981

Prepared for the National Center for Education Statistics under contract OE-300-78-0208 with the U.S. Department of Education. Contractors undertaking such projects are encouraged to express freely their professional judgment. This report, therefore, does not necessarily represent positions on policies of the Government, and no official endorsement should be inferred. This report is released as received from the contractor.

NCES 82-219

BLANK PAGE
DO NOT FILM

TABLE OF CONTENTS

PREFACE vii

ACKNOWLEDGMENTS xi

SUMMARY OF MAJOR FINDINGS xiii

1. INTRODUCTION: HISPANICS AND THE HIGH SCHOOL AND BEYOND SURVEY 1

1.1. Purpose of this Report 1

1.2. Background and Data Base 3

1.3. Definition of Hispanics and Hispanic Subgroups 5

1.4. Statistical Considerations 11

1.5. Significance Testing and Organization of the Tables 15

2. A PROFILE OF HISPANIC SUBGROUPS 18

2.1. Achievement 19

2.1.1. School Delay 19

2.1.2. Aspirations 24

2.1.3. Cognitive Achievement 27

2.2. Family Background 30

2.2.1. Father's Education 31

2.2.2. Family Income 33

2.3. Immigration History 34

2.3.1. Nativity 35

2.3.2. Length of Residence of Respondent 36

2.3.3. Length of Residence of Mother 38

2.4. Language 41

2.4.1. Home Language 42

2.4.2. Mother Tongue 45

2.4.3. Spanish Proficiency 50

2.4.4. English Proficiency 53

2.4.5. Spanish Usage..... 54

2.5. Conclusions..... 55

3. A BASELINE MODEL OF ACHIEVEMENT..... 57

3.1. Variables..... 59

3.2. Analysis..... 63

3.3. School Delay..... 65

3.4. Aspirations..... 75

3.5. Test Scores..... 77

3.6. Residual Correlations..... 79

3.7. Summary and Conclusion..... 82

APPENDIXES

A. Selected Questions on Language Use, Language Proficiency, and National Origin

B. Supplementary Tables: Age Distributions by Sex

REFERENCES

LIST OF TABLES

Table		Page
1.1.	Sample size and estimated population size by population subgroup: Spring 1980	6
1.2.	Percent distribution of language background by population subgroup: Spring 1980	8
1.3.	Percent distribution of self-reported race by population subgroup: Spring 1980	12
1.4.	Sample correction factors for percentages and means by grade and population subclass	16
2.1	Percent distribution of age by population subgroup: Spring 1980	22
2.2.	Percent distribution of years of schooling respondents expect to complete by population subgroup: Spring 1980	25
2.3.	Cognitive achievement scores on mathematics, reading, and vocabulary tests by population subgroup: Spring 1980	29
2.4.	Percent distribution of father's education by subgroup: Spring 1980	32
2.5.	Percent distribution of yearly family income by population subgroup: Spring 1980	34
2.6.	Percent distribution of nativity by population subgroup: Spring 1980	36
2.7.	Percent distribution of length of U.S. residence by population subgroup--how much of the student's life spent in the U.S.: Spring 1980	38
2.8.	Percent distribution of length of U.S. residence of mother by population subgroup--how much of mother's life spent in U.S.: Spring 1980	40
2.9.	Percent distribution of language usage at home by population subgroup: Spring 1980	44
2.10.	Percent distribution of respondent's pre-school language usage by population subgroup: Spring 1980	46
2.11.	Percent distribution of mother tongue type by home language type: Spring 1980	48

2.12. Spanish proficiency--composite scale of self-assessed ability to understand, speak, read, and write Spanish, by population subgroup: Spring 1980..... 52

2.13. English proficiency--composite scale of self-assessed ability to understand, speak, read, and write English, by population subgroup: Spring 1980..... 53

2.14. Composite scale of four indicators of Spanish use, by population subgroup: Spring 1980..... 55

3.1. Correlations, means, and standard deviations for models of scholastic achievement for Hispanics: Spring 1980..... 58

3.2. Coefficients of models of scholastic achievement for Hispanic sophomores: Spring 1980..... 66-67

3.3. Coefficients of models of scholastic achievement for Hispanic seniors: Spring 1980..... 68-69

3.4. Correlations among the errors for the model of a scholastic achievement: Spring 1980..... 80

B.1. Percent distribution of age by population subgroup for males: Spring 1980.....

B.2. Percent distribution of age by population subgroup for females: Spring 1980.....

PREFACE

The data and analyses presented in this report are from the first (1980) wave of the National Center for Education Statistics study, High School and Beyond, a longitudinal study of U.S. high school seniors and sophomores. This study was conducted for NCES by the National Opinion Research Center at the University of Chicago.

A detailed report on sample design and sampling errors, High School and Beyond: Sample Design Report, is available, so the sample will be described only briefly here. The sample was a two-stage stratified probability sample with schools within a stratum drawn with a probability proportional to their size. Once a school was selected, up to 36 sophomores and 36 seniors were drawn randomly from the students enrolled in each selected school.

Several special strata were included in the sample design. Schools in these special strata were selected with probabilities higher than those for schools in regular strata to allow for special study of certain types of schools or students. The following kinds of schools were oversampled:

- Public schools with high proportions of Hispanic (Cuban, Puerto Rican, and Mexican) students.
- Catholic schools with high proportions of minority group students.
- Public alternative schools.
- Private schools with high proportions of National Merit Scholarship finalists.

Substitutions were made for noncooperating schools in those strata where it was possible. Out of 1,122 possible schools, students at 1,015 schools and school administrators from 988 schools filled out questionnaires.

In many schools the actual number of seniors and sophomores was less than the target number for several reasons. First, in some schools fewer than

the number 36 sophomores or 36 seniors were enrolled. This reduced the number of eligible students from 73,080 (72 students in each of 1,015 schools) to 69,662. Second, 8,278 students were absent on the survey date. Third, 1,982 students, or in some cases their parents, declined to participate, exercising their right in a voluntary survey. Substitutions were not made for non-cooperating students. Finally, 1,132 cases were deleted because they contained only very incomplete information. Thus, data are available for 30,030 sophomores and 28,240 seniors. This represents a completion rate of 84 percent: 58,270 out of the 69,662 eligible students. In addition to the students in the regular sample, data were collected from friends and twins of participating students.

Weights were calculated to reflect differential probabilities of sample selection and to adjust for nonresponse. Using appropriate weights yields estimates for high school sophomores and seniors in the United States and separate estimates for schools or students classified in various ways, such as by geographical region or school type.

Information of several sorts was obtained in the survey. Students completed questionnaires of about one hour in length, and took a battery of tests with a total testing time of about one and one-half hours. School officials completed questionnaires covering items of information about the schools. A sample of parents of sophomores and seniors (about 3600 for each cohort) was surveyed primarily for information on financing of postsecondary education. Finally, teachers gave their perceptions of specified characteristics of students in the sample whom they had had in class, to provide information beyond the students' own reports about themselves.

This report is one of several analyzing High School and Beyond base year survey data. The study was designed to be relevant both to many policy

issues and to many fundamental questions concerning youth development and educational institutions. It is intended to be analyzed by a wide range of users, from those with immediate policy concerns to those with interests in more fundamental or long-range questions.

As succeeding waves of data on a subsample of these students become available (at approximately two-year intervals), the richness of the dataset, and the scope of questions that can be studied through it, will expand. In addition, use of the data in conjunction with NCES's study of the cohort of 1972 seniors (also available from NCES), for which data at five time points are now available, enriches the set of questions that can be studied.

The data are available on computer tape for a nominal fee from:

Statistical Information Office
National Center for Education Statistics
1001 Presidential Building
400 Maryland Avenue, SW
Washington, D.C. 20202
Phone: (202) 436-7900

BLANK PAGE
DO NOT FILM

ACKNOWLEDGEMENTS

The initial design of HIGH SCHOOL AND BEYOND was developed by the Longitudinal Studies Branch of the National Center for Education Statistics. Edith M. Huddleston, NORC project officer for HIGH SCHOOL AND BEYOND, and William B. Fetters, mathematical statistician, have guided this project since its inception, and have been responsible for many aspects of the research design. The current NCES project officer is Samuel S. Peng.

A study of this scope and magnitude would not have been possible without the active cooperation of many persons at various levels of educational administration: Chief State School Officers, Catholic Archdioceses and other private school organizations, principals and teachers in the schools, and of course, the students and their parents. The expertise, support, and persuasiveness of numerous study coordinators at participating schools was especially critical to the successful conduct of the study. Those who will use these data for the study of American education are deeply indebted to all these people.

A second debt is owed to all those people on the field and project staff of HIGH SCHOOL AND BEYOND, whose efforts brought into being the data that will make possible the study of issues involving young people and their schools, data on which the present report is based.

The funds for what was originally called the "Hispanic supplement" to High School and Beyond, i.e., the additional Hispanic students included in the sample and the extensions of the questionnaire on which much of this report is based, were provided by the Office of Bilingual Education and Minority Language Affairs (OBEMLA) and the Office for Civil Rights (OCR).

Special thanks are due to members of the National Planning Committee, who have been active in advising NCES on the design, implementation, and uses

of the study: Ellis B. Page, Chairman (Duke University), Robert F. Boruch (Northwestern University), Bruce K. Eckland (University of North Carolina, Chapel Hill), Barbara Heyns (New York University), David S. Mundel (Employment and Economic Policy Administration, City of Boston), Robert C. Nichols (State University of New York, Buffalo), Sally B. Pancrazio (Illinois Office of Education), and David E. Wiley (Northwestern University).

The National Opinion Research Center (NORC), under the direction of NCES, took responsibility for the remainder of the design and conducted the base-year survey; NORC's preliminary analysis of the baseyear data contributed to the development of this publication. James S. Coleman served as Principal Investigator at NORC, with Carol B. Stocking as Project Director. Other contributing NORC staff members were Fansayde Calloway, who directed field work for the project, and Antoinette Delk, Larry Dornacker, Martin Frankel, and Natalie Suter.

Finally, the obvious contributions of editors Susan Campbell and Suzanne Erfurth and the Data and Word Processing staffs at NORC are appreciated. We also acknowledge the work of Robert Aponte, who prepared the tables for the report.

SUMMARY OF MAJOR FINDINGS

(Definitions of terms and exact figures are presented in later sections of the report.)

Demographic Characteristics

. Proportions of Hispanic students in the tenth and twelfth grades with Spanish language background vary markedly by Hispanic subgroup, from about 80 percent for Cuban seniors to about 38 percent for Other Latin American seniors (Table 1.2).

. The proportion of students whose fathers have a college degree is higher for Cubans and Other Latin Americans than for other Hispanic subgroups. The proportion for each of these two groups is also higher than that for blacks but considerably lower than that for whites not of Hispanic origin. The proportion of students whose fathers have not completed high school is also higher for most Hispanic subgroups than it is for whites and blacks (Table 2.4).

. Puerto Rican families appear to be the most disadvantaged with respect to income, while Cubans and Other Latin Americans appear to be the most advantaged. All Hispanic subgroups, except Puerto Ricans, do better than blacks in this respect. All Hispanic groups fare less well than non-Hispanic whites. For example, in comparison with non-Hispanic whites, twice as many Cuban seniors and almost five times as many Puerto Rican seniors reported family incomes of less than $12,000 per year (Table 2.5).

. Hispanic subgroups differ greatly with respect to immigration history. Mexican-American and Other Latin American families of the 10th and 12th graders have been settled in the United States for the longest period of time. The families of Cuban students have entered the country most recently. Puerto Ricans are in between Mexican-Americans and Other Latin Americans on the one hand and Cubans on the other with respect to the length of settlement of the family in the United States. This conclusion is supported by the distribution of places of birth of students, their length of residence in the country, and the length of residence of their mother (Table 2.7).

Language Use

. Use of the Spanish language by these students varies among Hispanic subgroups. Considering several indicators together (home language use, mother tongue, self-assessed proficiency in Spanish, frequency of use of Spanish with parents), Other Latin American and Mexican-American students appear to be the most linguistically assimilated subgroups, Cubans the least assimilated, and Puerto Ricans in between (Tables 2.9 and 2.10).

. However, subgroups are very similar with respect to scores on the self-assessed English proficiency measure, with Cuban students reporting a slightly higher proficiency (Table 2.13)

Academic Achievement

. Rates of school delay (being two or more years older than the modal age for a grade) are considerably larger for Hispanic seniors than for white seniors not of Hispanic origin (Table 2.1).

. Hispanics generally have lower educational aspirations than either blacks or whites, when the level of aspiration is calculated as the percentage of a group that expects to achieve at least a college degree. Cubans are an exception, with the highest aspiration level of all groups (Table 2.2).

. Average scores on mathematics, reading, and vocabulary tests for Hispanic subgroups are intermediate between those of blacks and non-Hispanic whites. Among Hispanics, Cubans have the highest scores on all three tests, even though this group makes more use of the Spanish language (Table 2.3).

. In multivariate analyses of several measures of educational achievement for Hispanics (school delay, educational aspirations, scores on mathematics, reading, and vocabulary tests) a number of factors are generally found to have substantial effects. In addition to socioeconomic status of the family, proficiency in English and proficiency in Spanish are positively related to achievement. Perhaps surprisingly, the length of residence of the family in the United States is negatively related to achievement, and so is the frequency of use of the Spanish language (see Chapter 3).

. Average differences in achievement among Hispanic subgroups generally disappear when individual factors (such as socioeconomic status, sex, linguistic indicators, immigration history) are controlled for, except for the difference between Cubans and other Hispanics: Cuban 10th and 12th graders still achieve better in school than other groups (see Chapter 3).

CHAPTER 1

INTRODUCTION: HISPANICS AND THE HIGH SCHOOL AND BEYOND SURVEY

1.1. Purpose of This Report

The fate of Hispanics in U.S. schools has been the subject of increasingly frequent policy debates (Weinberg, 1977). Most policy issues--such as the desirability of bilingual-bicultural education, the effects of segregation, tracking, and so forth--concern the process of education: the effects of school behavior, or some school characteristic, on the outcome of education for Hispanics. Much research on these issues has been inconclusive for lack of a suitable data base, either with respect to the composition of the sample or with respect to the completeness of the information gathered. High School and Beyond (HS&B) will contribute toward filling this gap by providing both sufficient samples of Hispanics from various origins, and detailed information on a variety of policy-relevant factors, including the linguistic practices of students, their exposure to bilingual education, their immigration histories, and the degree of segregation in their schools and in their communities. There is no doubt that the data will be extensively explored by researchers interested in these process issues.

The main goal of this report is to pave the way for such analyses of the process of education. To do this, we focus on the preliminary task of assessing the effects associated with the input to education: the composition of the Hispanic student population in the 10th and 12th grades with respect to various characteristics such

as language use and proficiency, length of contact with U.S. society, the socioeconomic status of the family. Any refined analysis of the process of education, such as evaluations of the effect of exposure to bilingual education, would have to control for such input factors.

We have therefore organized the substantive material of this report into two parts. First, in chapter 2, we attempt to provide a basic description of the differences among Hispanic subgroups, and between Hispanics and the rest of the student population, with respect to important aspects of the achievement process. The factors examined include both measures of achievement (school delay, aspirations, test scores) and possible explanatory factors (language usage, socioeconomic status, immigration history). This section is deliberately descriptive. Then, in chapter 3, we present a basic model relating selected input factors to various measures of the achievement of Hispanics in U.S. schools. As we emphasize later, the HS&B data enable us to investigate characteristics associated specifically with the Spanish heritage of the students, such as linguistic indicators, in addition to factors of a more general nature, such as the socio-economic status of the family. Our purpose there is to assess how much of the variation in achievement by Hispanics can be explained solely on the basis of these input factors, before introducing characteristics of the process of education. Prior to the presentation of the results, we provide in the remainder of this chapter a discussion of the High School and Beyond data, the selection of the samples, and other methodological aspects of the study.

1.2. Background and Data Base

With respect to Hispanics, the High School and Beyond data are invaluable because of two characteristics that are rarely found in combination: the variety of the specially relevant information collected, and the size of the samples of Hispanics from different national origins.

First, with respect to the data, High School and Beyond is the second of the only two large-scale studies conducted among students in the United States that include detailed questions on language use and language proficiency, in addition to the general information usually gathered in surveys of that nature. (The other study providing refined information on language use is the Survey of Income and Education, conducted in 1976 in a collaboration between the Bureau of the Census and NCES. However, the language questions in the High School and Beyond survey are more detailed than those in SIE.) Furthermore, High School and Beyond has gathered additional information on the nativity of the respondents, their length of residence in the country, the length of residence of both parents, several measures of contact of respondents with some form of bilingual education, and the history of segregation in the respondents' school (at current time and at grade levels 1, 6, and 9). This probably makes the High School and Beyond data base the most complete ever compiled with respect to Hispanics (as well as other language minorities). Tables presenting this information are included throughout this report.)

Second, the oversampling of Hispanics has produced a total of 6,698 Hispanic students, either sophomores or seniors. Among them,

Cubans were oversampled most heavily so as to gather a sufficient number of cases from that group, which constitutes a small fraction of the general population. The exact numbers in each subgroup are presented and discussed in the next section. Furthermore, the sampling design allows computation of weighted figures that are representative of the population as a whole. Further discussion of this topic is presented in section 1.4.

It should be clear that the design of High School and Beyond has particular advantages for the study of Hispanic sophomores and seniors in the United States. This study combines the advantages and avoids the defects of the two categories of studies that have traditionally been conducted. On the one hand, it contains the specific information that was previously available only in very small surveys (e.g. a sample of students in a border town in Texas, or a sample from a Puerto Rican neighborhood in New York). On the other hand, it is a nationally representative study that supplements the general information usually collected (e.g. about family background, aspirations, attitudes) with information that is especially relevant for language and cultural minorities..

As succeeding waves of data on a subsample of these students become available (at approximately two-year intervals), the richness of the dataset, and the scope of questions that can be studied with it, will expand. In addition, use of the data in conjunction with NCES's study of the cohort of 1972 seniors (also available from NCES). for which data at five time points are now available, will enrich the

set of questions that can be studied (though the sample of Hispanics in the 1972 NLS was not as large, and the information concerning language minorities very incomplete).

In the next two sections of this introductory chapter we discuss a variety of methodological considerations with respect to the data on Hispanics: the definition of Hispanics and national origin subgroups, and statistical considerations concerning the samples.

1.3. Definition of Hispanics and Hispanic Subgroups

For this report, the classification of Hispanics was based on respondents' answers to the following question: "What is your origin or descent? (If more than one, please mark below the one you consider the most important part of your background.)" (See Appendix A for the full text of the question.) Under the general heading of "Hispanic or Spanish" were grouped four possible answers: (1) Mexican, Mexican-American, Chicano; (2) Cuban, Cubano; (3) Puerto Rican, Puertorriqueno or Boricua; (4) Other Latin American, Latino, Hispanic, or Spanish descent.

From the answers, we constituted four groups in the obvious ways, labelled, for simplicity: "Mexican-American," "Cuban," "Puerto Rican," "Other Latin American." Table 1.1 presents the weighted and unweighted numbers of Hispanic sophomores and seniors belonging to each group classified by this procedure. These samples constitute the basis of the tables and analyses presented in the report. However, in many cases the usable totals are smaller because of missing answers.

Table 1.1.--Sample size and estimated population size by population subgroup: Spring 1980

Subgroup	Sample size 1/	Population size 2/
Sophomores		
Mexican-American	2,123	149,780
Cuban	306	16,025
Puerto Rican	369	41,625
Other Latin American	723	75,776
Seniors		
Mexican-American	1,893	102,477
Cuban	334	11,223
Puerto Rican	308	18,145
Other Latin American	642	55,810

1/ Actual (unweighted) number of respondents; unadjusted for probability of selection.

2/ Weighted estimate of population size; adjusted by probability of selection.

The question used to determine whether a student is Hispanic is based on self-identification by the respondent. This is in agreement with the emerging consensus among researchers on what constitutes "ethnic" identity, and this type of question has been shown to be the most efficient in eliciting a positive national origin identification from respondents in the general population (Smith, 1980). However, it is important to note for comparison purposes that a variety of alternative criteria have historically been used to define the Hispanic population (Hernandez et al., 1973). Smith classifies the various methods into three broad categories: natal definitions, based on the country of birth of the respondent, or of the parents; behavioral

definitions, based on some objective cultural criterion such as the use of a language other than English; and subjective criteria, involving self-identification by the respondent. The definition of Hispanics used in this report is an example of the subjective type of definition.

One potential problem in comparing the figures presented in this report with other bodies of data, such as those produced by the Census, is that the populations defined by alternative criteria may overlap only very imperfectly. Many tables presented in chapter 2 of the report illustrate the lack of fit between the subjective definition and several alternative ones, but to give an idea of the potential magnitude of the discrepancy, we present in table 1.2 the relationship between the subjective definition used here and a very broad behavioral (or objective) definition of Hispanics: Spanish language background. A respondent was considered as being of Spanish language background if "Spanish" was answered to at least one of five language questions: mother tongue of respondent (first language spoken), "second mother tongue" (other language spoken before schooling), usual language spoken at home, other language spoken at home, and usual language of respondent.

This is a very broad definition; it includes even respondents who were raised as English monolinguals but whose parents occasionally speak Spanish at home, and respondents from homes where Spanish is no longer spoken, whose mother tongue is English, but who learned to speak some Spanish when they were children. But table 1.2 shows that using the broad behavioral definition of a Hispanic as someone of Spanish language background would fail to identify quite substantial fractions

Table 1.2.--Percent distribution of language background by population subgroups: Spring 1980

Subgroup	Sample size	Total	Spanish background 1/	No Spanish background
Sophomores				
Mexican-American	2,123	100.0	66.6	33.4
Cuban.....	306	100.0	75.0	25.0
Puerto Rican	369	100.0	72.9	27.1
Other Latin American ...	723	100.0	31.5	68.5
Seniors				
Mexican-American	1,893	100.0	72.0	28.0
Cuban	334	100.0	79.6	20.4
Puerto Rican	308	100.0	79.6	20.4
Other Latin American ...	642	100.0	37.5	62.5

NOTE: Percentages are weighted.

1/ Respondent answered "Spanish" on at least one of five language questions concerning first language spoken; other pre-school language usage; usual language spoken at home; other language spoken at home; and usual language of respondent.

of respondents who identified themselves as Hispanics: for sophomores, from 25 percent for Cubans to 68.5 percent for Other Latin Americans. The high proportions of Other Latin Americans with no Spanish language background are striking and suggest a population that is largely assimilated linguistically. In view of this finding, it is appropriate to comment briefly on the meaning of the Other Latin American category.

In view of its residual nature, one would expect the Other Latin American category to be quite heterogeneous. A respondent could have checked that category for a variety of reasons, among which the following are possibilities:

1. The respondent's national origin was simply not among those listed in the question. For example, a Dominican student would have to have answered in that way. Both respondents who are recent immigrants and respondents with families long established in the United States may have answered Other Latin American for that reason. There is no necessary relationship between the propensity to answer that way and immigration history.

2. A respondent may have been conscious of being Hispanic, but simply not have known the national origin of the family precisely. In such cases, the lack of knowledge is likely to have been related to the distance from contact with the country of origin and, as Smith (1980) argues, to the length of settlement in the United States.

3. Despite the instruction in the question to mark "the most important" part of the background in case of several national origins, a respondent with mixed background (e.g., Cuban father and Puerto Rican mother) may have been unwilling to choose between the two origins and answered Other Latin American as a compromise. As Smith (1980) also argues, the likelihood of mixed background is more likely for populations long established in the country. This mechanism is what causes most difficulties in eliciting a positive national origin identification in the general population.

If mechanisms of type (2) and (3) are important in explaining Other Latin American answers, one would expect a large part of the group to be composed of children of earlier immigrants, largely assimilated into English. The large proportions of Other Latin Americans with no Spanish language background is consistent with this possibility. Further evidence is presented in chapter 2. However, the group may also include a substantial fraction of recent immigrants because of type (1) mechanisms, increasing the heterogeneity of the group. Evidence that this is indeed the case is also presented later in the report.

The reader should also be aware of another issue concerning the "origin or descent" question. A large number of respondents chose to ignore the "most important" clause of the question and marked two

or more national origins. These multiple answers are coded with a special missing value in the file. Among sophomores, 2,965 such answers were recorded, and 2,297 among seniors, corresponding to 9.9 and 8.1 percent, respectively, of the samples for the two cohorts. It may be that a substantial fraction of these multiple-punch answers include a Hispanic subgroup as one of the choices, and a researcher might want to include these students in the Hispanic sample. However, such a strategy would involve conceptual as well as technical decisions. While it would seem reasonable to consider a student marking two different Hispanic origins as Hispanic, it is less clear what should be done with a student listing, say, both Polish and Mexican-American. The multiple answers to the origin or descent question certainly provide room for further analyses by researchers willing to make the necessary substantive choices.

To complete this discussion of the various population subgroups that are distinguished in later analyses, we describe two additional samples that are used. For comparison purposes, and to reduce computation costs, we selected from the main file random samples of 1,000 whites not of Hispanic origin and 1,000 blacks not of Hispanic origin. Each case in the file had an equal chance of being selected. A respondent could be included in these samples only if he/she had reported both a non-Hispanic origin and race as black or white. These comparison samples are therefore based on a more restrictive definition of "black" and "white" than is often used in published tables based on other bodies of data. This is important since there is a tendency to assume that

blacks and Hispanics constitute clearly distinguishable minorities. However, especially for Hispanics, the relationship between the concept of race and the concept of national origin or ethnicity may be quite problematic.

To illustrate this, table 1.3 presents the distribution of answers to the question about race ("What is your race?") for the four Hispanic subgroups. The table reveals very different perceptions of race across subgroups: while most Cubans and Other Latin Americans consider themselves white (76.3 and 67 percent, respectively, for sophomores), the most frequent answer is "other" for Mexican Americans and Puerto Ricans (47.2 and 57.4 percent, respectively). The high percentages of "other" answers are not surprising in the Hispanic context, given the lack of clear distinction between the concept of race and national origin (as examplified in the use of the Spanish term la raza to designate persons of Hispanic origin). Another noteworthy feature of the data in the table is the nonnegligible fraction of all groups reporting race as black (from 7.1 to 13 percent for sophomores). This constitutes a substantial overlap between two supposedly distinct minorities. However, it should be kept in mind that race here is self-reported. Results might be different if, for example, race were imputed by an interviewer on the basis of physical features of the respondent.

1.4. Statistical Considerations

Two aspects of the representativeness of the results presented in this report deserve further comment: the procedure of weighting the data, especially for the Cuban subsample, and the degree to which

Table 1.3.--Percent distribution of self-reported race by population subgroup: Spring 1980

Subgroup	Sample size	Total	Black	White	American Indian/ Alaskan Native	Asian	Other
Sophomores							
Mexican-American	2,092	100.0	7.1	43.0	2.6	.1	47.2
Cuban	306	100.0	7.4	76.3	3.3	1.5	11.5
Puerto Rican	365	100.0	11.0	28.1	2.1	1.4	57.4
Other Latin American ...	716	100.0	13.0	67.0	.8	1.0	18.3
Seniors							
Mexican-American	1,869	100.0	7.6	38.5	1.7	.1	52.1
Cuban	329	100.0	10.0	77.6	0.0	3.0	9.4
Puerto Rican	307	100.0	7.8	29.8	1.6	1.5	59.4
Other Latin American ...	632	100.0	13.5	65.2	.3	0.0	20.9

NOTE: Percentages are weighted.

the samples provide information on the entire Hispanic population in the corresponding age groups living in the United States, as opposed to the population of Hispanics who have remained in school at grades 10 and 12.

As mentioned earlier, schools with special characteristics were oversampled to allow for special studies. Schools with large proportions of Hispanics were among those selected with greater probability. Schools with large numbers of Cubans were included with even greater probabilities. (These schools are all in Florida or New Jersey, except for one located in California.) As a consequence, data for Cubans in these schools were assigned a very small weight in analyses presented in this report. However, a number of Cubans were found in schools of the normal strata, selected with low probability and usually located outside the areas of greater Cuban concentration. As a result, Cubans in these schools were assigned weights considerably greater than the weights of Cubans in the Florida and New Jersey areas of concentration.[1/] Insofar as Cuban students found outside Cuban enclaves might be expected to have somewhat different characteristics as compared to those in Florida and New Jersey, then their contribution in computing the weighted figures is considerably inflated: a few atypical cases might unduly influence the results.

To check for this possibility, we also computed the tables of chapter 2 without weighting and compared them with the weighted figures. In most cases there were not substantial differences between weighted and unweighted figures for Cubans. In one instance where a large

---

1/ For the Cubans in the oversampled schools, the sample sizes are 256 and 216 and the average weights are 17.0 and 24.9 for seniors and sophomores respectively. For the rest of the High School and Beyond sample, the numbers of Cubans are 78 and 90 with average weights of 56.7 and 91.2 for seniors and sophomores respectively.

discrepancy was found, we note it in the text and report the unweighted figures for comparison purposes.

The second issue is the representativeness of the sophomore and senior cohorts for the entire Hispanic population of corresponding age. There is considerable evidence that dropout rates are substantially greater for Hispanics than for the rest of the population. Estimates based on data from the Survey of Income and Education of 1976 (1978, table 1) indicate that among those aged 14 to 25 about 24 percent of Hispanics and only 10 percent of non-Hispanics had not graduated from high school and were not enrolled in school.[1/] Among Hispanic dropouts, 60 percent had left school before grade 10, while the remaining 40 percent never completed grade 12 (NCES, n.d.). Therefore, for both the sophomore and senior cohorts, Hispanics represent a more "select" group than the rest of the population. Since teenagers who had left school before the survey was taken are presumably the most marginal with respect to many characteristics of family background and measures of achievement, estimates of the relative disadvantage of Hispanics based on the High School and Beyond data probably underestimate the disadvantage in the population of corresponding age. Therefore, generalization of the results to that population should only be made with extreme caution. Further methodological consequences of the greater selectivity of the Hispanic sample are discussed in Nielsen (1980). Note, however, that the dropout rates estimated from the 1976 SIE should be considered as indicative only in this context, since students in the High School and Beyond sample belong to a younger cohort, and there is evidence that school nonattendance patterns for Hispanics have become increasingly

---

1/ With cross-sectional data, such as the S.I.E., the dropout rate is estimated as the overall percentage of persons aged 14-25 who are not attending high school and have not graduated. Note that this estimate might be misleading, in one direction or another, for a variety of reasons.

similar to those for the population as a whole in recent years (U.S. Commission on Civil Rights, 1978, p. 10; see, however, NCES, 1980, pp. 96-97). When data from the High School and Beyond followup survey in 1982 become available, more refined estimates of the rates of leaving school between grades 10 and 12 will be possible.

1.5. Significance Testing and Organization of the Tables

All the analyses contained in chapters 1 and 2 of this report have been performed on weighted data in order to provide population estimates. Aside from the cautions expressed in section 1.4. with regard to Cubans, it should be noted that the weighting implies that one cannot infer the sample size in a particular cell of any table from the information reported. The tables show unweighted sample sizes and percentages based on weighted numbers; therefore, the sample size cannot be used as a base with which to compute the actual number of students in any cell. Any such inference may be quite erroneous due to the fact that different students have different weights.

Standard errors or confidence intervals are not reported in the tables for chapters 1 and 2.[1/] However, the tables in chapter 1 and 2 and the information provided in this section allow the calculation of approximate standard errors for percentages and means.

The general equation for calculating the approximate standard error of a percentage is:

$$s.e.(p) = D\sqrt{p(100-p)/n}$$

1/ T-values are reported for the regression coefficients in chapter 3, however, these have been computed assuming simple random sampling and have not been corrected to account for the HS&B sample design effect. Since clustered samples tend to produce larger variances, the t-values reported for the regressions in chapter 3 are slightly non-conservative. Those researchers who want to correct for the non-conservative bias of the sample design might adopt a more conservative probability level (e.g., .03 as opposed to .05) in assessing statistical significance.

Table 1.4--Sample correction factors for percentages and means by grade and population subclass

Population subclass	Sophomores		Seniors	
	D [1]	DE [2]	D [1]	DE [2]
Whites	1.4	1.3	1.5	1.3
Blacks	1.5	1.4	1.4	1.4
Hispanics	1.5	1.3	1.5	1.3
Males	1.5	1.3	1.4	1.3
Females	1.4	1.4	1.4	1.4

1/ D is the correction factor for percents.

2/ DE is the correction factor for means.

where p is the percentage for which the standard error is to be calculated; s.e.(p) is the approximate standard error of p; D is a correction factor, which increases with the departure of the sample from a simple random sample as a result of clustering or other aspects of sample design; and n is the unweighted number of students in the particular class over which the percentage is calculated.

One can compute approximate standard errors for means as follows:

$$se(\bar{x}) = DE\sqrt{\frac{S^2}{n}}$$

$se(\bar{x})$ is the approximate standard error of the mean; S^2 is the weighted variance estimated for the demographic subclass and grade cohort from which the mean was computed; DE is a factor that corrects for the effect of the sample design; n is the unweighted sample size for the particular mean.

The values of D and DE for the classes relevant to this report are given in table 1.4. Values for n, p, and $\bar{x}$ can be found at the appropriate table in the text. When percentages or means are based on other classifications or on subclassifications within each of these groups, it is appropriate to use the subclass size together with the largest correction factor of those shown in the table that could apply to the class or subclass in question.

CHAPTER 2

A PROFILE OF HISPANIC SUBGROUPS

In this chapter we provide a general description of the composition of the Hispanic student population in grades 10 and 12 with respect to four broad categories of factors. We examine selected indicators of educational achievement, of the socioeconomic status of families, of immigration history, and of language use and proficiency. The purpose of this examination is to provide an overall picture of the major Hispanic subgroups (Mexican-Americans, Cubans, Puerto Ricans, Other Latin Americans) as they compare with "Anglos," or whites not of Hispanic origin, and blacks of a non-Hispanic background.

The usefulness of a descriptive presentation of this nature is twofold. First, it provides basic information that is often not available elsewhere for these grade cohorts, such as data on language use and proficiency. Second, Hispanic subgroups will be shown to differ considerably with respect to educational achievement and to factors likely to influence achievement, such as family background, immigration history, and linguistic habits. The systematic differences between groups suggest possible mechanisms relating these factors to achievement, and these findings lead naturally to the more elaborate causal analyses of chapter 3.

To derive the tables in this chapter, we used the four subsamples of Hispanics and the samples of blacks and whites described in chapter 1. We include information on blacks and whites only when the comparison

with Hispanics would be meaningful; for the linguistic indicators, for example, we consider only Hispanics. The figures are weighted to represent population estimates. To reduce the length of the discussion we also comment only on seniors when the pattern for sophomores is the same. We choose to focus on seniors because they are at a stage of their educational careers that is both closer to an important transition (graduation from high school with the choice of either going to college or joining the labor force) and more representative of the total impact of schooling.

2.1. Achievement

The comparatively low achievement of Hispanics in school is central in policy debates. This is the case not only because some aspects of achievement seem desirable in themselves, such as the acquisition of useful skills and knowledge, but also because education is a prerequisite to so many prestigious, lucrative, or otherwise desirable occupations. For the individual, low educational achievement virtually guarantees low occupational achievement. Among all possible indicators of achievement, we selected three: school delay, aspirations and cognitive achievement.

2.1.1. School Delay

School delay is often viewed as a key factor in the low educational achievement of Hispanics (e.g., ASPIRA, 1976; Carter and Segura, 1979). Delay is certainly undesirable in itself, since it means a greater expenditure of resources (time, energy, money) to acquire a given

credential such as a high school degree. But, perhaps more important is the fact that delayed students may demonstrate a greater propensity to drop out of school entirely for a number of reasons: discouragement, disjuncture between the normal physical-psychological maturation of the delayed student and school conditions tailored for a younger cohort, increased financial demands on the older student, and the increasing attractiveness and availability of job opportunities outside the school.

School delay has become an important issue for Hispanics for two principal reasons. First, delay rates are markedly higher for Hispanics than for the rest of the population. Second, there is considerable geographical variation in the delay rates among Hispanics. Carter and Segura (1979) document the variation in delay rates across states for Mexican-Americans, and ASPIRA (1976) the variation across metropolitan areas for Puerto Ricans. The geographical differences might suggest that the incidence of grade repetition for Hispanics depends in large measure on the policies and practices of local school systems: for example, the delay rate of Mexican-Americans might be greater in Texas than in California because school authorities there are more prone to force a student to repeat a grade, as Carter and Segura suggest. The policy implication is that reforming school policies would reduce the proportion of delayed students and minimize the harmful consequences of being left behind However, the variation in delay across states and metropolitan areas might also be entirely due to individual causes. For example, given the composition of a student body, more students might be delayed in one place than another because their linguistic and socioeconomic handicap is more

severe there, even though school policies are exactly the same everywhere. If this is the case, changing school policies would have little effect on the desired outcomes. In chapter 3, we present evidence that individual causes, including measures of English proficiency, have very little explanatory power for delay compared with their impact on other measures of achievement. The implication is that variations in school policies and practices may be the most important factor. In this chapter, we concentrate on a comparison of delay rates across Hispanic subgroups, and between Hispanics and the rest of the population.

To capture variations in delay rates among population subgroups in the simplest way, table 2.1 presents the distribution of age by population subgroup for both sophomores and seniors, both sexes combined. (We discuss differences by sex later.) The table also reports the total percentages of students whose age is two or more years above the modal age for the grade in the population as a whole (15 and 17 years old for sophomores and seniors, respectively), a common definition of school delay.[1/] Among seniors, Mexican-Americans and Puerto Ricans have the highest rates of delay: 9.8 and 12.6 percent, respectively. These figures are considerably higher than those for whites (2.5 percent) and even blacks (7 percent). Cuban seniors have the lowest delay rate among Hispanics (6.4 percent).

As argued in chapter 1, figures for seniors may conceal the true extent of delayed schooling, as it affects a cohort over the entire

---

1/ Note, however, that this definition is not the same as another one commonly used with cross-sectional surveys: the percentage of an age-group two or more years behind the modal grade for that age-group.

Table 2.1.--Percent distribution of age by population subgroup: Spring 1980

Subgroup	Sample size	Total	13 or younger	14	15	16	17	18	19	20	21 and older	At least 2 year delay 1/
Sophomores												
Mexican-American	1,926	100.0	0.3	0.8	40.8	45.0	10.8	1.5	0.4	0.1	0.2	13.0
Cuban	292	100.0	--	1.1	41.3	43.1	13.2	0.1	1.2	--	--	14.5
Puerto Rican	341	100.0	0.6	2.0	43.0	42.9	10.0	1.5	--	--	--	11.5
Other Latin American	652	100.0	0.5	0.4	43.2	46.4	7.9	1.2	0.2	--	0.0	9.3
Non-Hispanic black	878	100.0	0.8	0.5	43.3	42.6	10.3	1.9	0.2	0.1	0.4	12.9
Non-Hispanic white	964	100.0	--	0.3	50.4	44.4	4.5	0.3	0.1	--	--	4.9
Seniors												
Mexican-American	1,810	100.0	--	--	0.1	1.0	44.2	44.9	8.4	1.0	0.4	9.8
Cuban	330	100.0	--	--	--	1.5	48.5	43.6	5.2	0.3	1.0	6.4
Puerto Rican	293	100.0	--	--	0.9	1.9	40.7	44.0	10.0	1.4	1.2	12.6
Other Latin American	586	100.0	--	--	0.4	1.3	45.0	44.3	7.4	1.2	0.2	8.8
Non-Hispanic black	908	100.0	--	--	0.4	1.3	47.9	43.4	5.3	1.2	0.5	7.0
Non-Hispanic white	974	100.0	--	--	--	0.9	53.4	43.2	2.2	0.2	0.1	2.5

NOTE: Percentages are weighted.

1/ Total percent of students whose age is at least two years above the modal age for the grade in the population as a whole (sophomore modal age = 15; senior modal age = 17). This column is not included in the percent distribution.

educational career, since at that stage many of the delayed students may have already dropped out of school. Using the 1976 Survey of Income and Education figures discussed in chapter 1, one would estimate that 60 percent (the proportion of Hispanic dropouts leaving school between grades 10 and 12) of 24 percent (the total dropout rate for Hispanics), or roughly 14 percent of all Hispanics, drop out between grades 10 and 12. If one assumes in addition that the rate of "recovery" from delay (skipping a grade after being delayed one) is negligible, and that most instances of grade repetition occur before the sophomore year, one would expect delay rates to be greater for sophomores if students left behind have a greater propensity to leave school than those on the normal schedule. A comparison of delay rates for seniors and sophomores in table 2.1 reveals that sophomores do indeed have larger delay rates for most groups. The differences are particularly strong for Mexican-Americans, Cubans and blacks. Puerto Ricans are the exception: the delay rate for seniors is greater than the rate for sophomores. One could speculate that this is due in part to a pattern of commuting migration between Puerto Rico and the mainland, with a substantial fraction of students being forced to repeat a grade between grades 10 and 12 because of the obvious difficulties of certification and adjustment associated with the change from one school system to another. Such a mechanism might compensate for the higher attrition rate for delayed students between the two grade levels, even though there is some evidence that it is very high for Puerto Ricans relative to other Hispanics (Gomez-Day, 1980).[1/] Data from the High School and Beyond followup survey in 1982 will undoubtedly shed more light on such processes.

1/ We are indebted to Rafael Valdivieso for a penetrating comment on this topic, and for pointing our attention to important recent sources of data.

Table 2.1 presents the age distributions for both sexes combined. However, it has been argued that male and female students differ substantially with respect to delay, and with respect to the incidence of and motivations for leaving school (Gomez-Day, 1980, based on the Department of Labor's National Longitudinal Study). To capture this, we discuss briefly the differences in age distributions between males and females. The relevant tables are contained in appendix B. For sophomores, there is a general pattern: male students have a higher incidence of delay than females for all groups. The sex differences are particularly strong for Mexican-Americans, Puerto Ricans, and blacks. For seniors, however, the pattern of sex differences is much less systematic: it is reversed for Puerto Ricans and Other Latin Americans (females are more likely to be delayed than males), and the sex difference is very much attenuated for Mexican-Americans. It is quite possible that these interactions among grade, sex, population subgroup, and delay are due to the interrelationship of individual and environmental mechanisms that differ among subgroups and by sex. However, it is difficult to speculate about them with any confidence at this stage since the sophomore and senior samples in the data file consist of different individuals. Again, further waves of the High School and Beyond survey should provide precious information on these issues.

2.1.2. Aspirations

We use educational aspirations here as a proxy for later educational achievement. Aspirations have been shown to be one of the best predictors of actual achievement (Otto and Haller, 1979). Table 2.2 presents the distribution of answers to the question: "As things stand now, how

Table 2.2.--Percent distribution of years of schooling respondents expect to complete by population subgroup: Spring 1980

Subgroup	Sample size	Total	Less than high school completion	At least high school, but less than four years college	4-year college degree	Master's degree	Ph.D., M.D. or other advanced degree	Total college 1/
Sophomores								
Mexican American	2,031	100.0	2.5	69.5	14.0	6.6	7.3	27.9
Cuban	292	100.0	1.7	48.3	22.6	6.6	20.9	50.1
Puerto Rican	354	100.0	2.1	62.0	17.4	8.0	10.5	35.9
Other Latin American.	691	100.0	1.6	61.5	21.5	6.1	9.2	36.8
Non-Hispanic black ..	939	100.0	1.7	56.5	23.1	6.9	11.7	41.7
Non-Hispanic white ..	971	100.0	1.0	55.6	25.0	8.7	9.7	43.4
Seniors								
Mexican American	1,857	100.0	1.1	65.3	19.0	8.6	6.0	33.6
Cuban	327	100.0	0.7	44.4	22.1	17.2	15.6	54.9
Puerto Rican	302	100.0	1.0	64.4	15.8	11.2	7.5	34.5
Other Latin American.	631	100.0	1.0	62.2	20.0	7.2	9.5	36.7
Non-Hispanic black ..	963	100.0	0.9	53.6	24.3	11.2	9.9	45.4
Non-Hispanic white ..	977	100.0	0.2	56.2	23.9	10.7	9.0	43.6

NOTE: Percentages are weighted.

1/ This column is not included in the percent distribution.

far in school do you think you will get?" To make the table easier to interpret, we regrouped several categories (two years of vocational education, two years of college, etc.) into one: at least high school but less than four years of college. We also regrouped under the category "total college" all respondents who answered either a college degree, a master's degree, a Ph.D., or other advanced degrees.

Considering the total-college category, strong differences emerge among Hispanic seniors. Cubans appear to have very high educational aspirations, even higher than those of non-Hispanic whites (54.9 percent versus 43.6 percent who think they will obtain at least a college degree). The remaining three Hispanic subgroups have aspirations levels that are similar to one another, with percentages ranging from 33.6 to 36.7 for seniors. Blacks have aspiration levels very similar to the ones for whites. With respect to Cubans, it should be noted that their over-representation in the total-college category respective to whites is entirely due to their overrepresentation in the two most advanced educational categories: M.A. or Ph.D. or equivalent. One might speculate that the high aspiration levels of Cubans are due to the professional elite component of the parents of these students. More evidence on this is presented later in this report.

Comparing seniors and sophomores, it appears that the same pattern of differences between groups holds for the two cohorts, except maybe for Mexican-Americans, among whom sophomores have lower aspiration levels than seniors (28 and 33.6 percent in the total-college category, respectively). The overall similarity is somewhat surprising, since one would expect seniors to have much more concrete ideas about the

maximum level of educational achievement they are able, and willing, to reach. The similarity suggests that ultimate educational goals may be formulated quite early during the student's career, before grade 10.[1/]

Another striking feature of table 2.2 is the uniformly small percentages in the less-than-high school-completion category, with the maximum of 2.5 percent for sophomore Mexican Americans. The figures are systematically higher for sophomores, presumably reflecting the fraction expecting to leave school between grades 10 and 12, but the small numbers are nowhere near the 14 percent estimate of the dropout rate of Hispanics between the two grades that can be computed from the SIE data. One hypothesis is that a large fraction of those dropouts complete grade 9 but never start grade 10, or leave school shortly after the beginning of the school year, thus missing the High School and Beyond survey in winter or early spring. Another hypothesis, which we find more plausible, is that the decision to leave school is typically not planned in advance by the student concerned: he/she leaves school when faced with adverse circumstances perhaps associated with remaining in school and acceptable alternatives to school; these temporarily combine to make continuation undesirable or impossible. The convergence of bad luck is probably rarely foreseen.

2.1.3. Cognitive Achievement

Cognitive achievement, as measured by tests such as the ones administered as part of High School and Beyond, is important in policy

---

1/ Researchers interested in studying the timing of educational aspirations further might want to analyze answers to question 68 of the senior questionnaire, in which respondents were asked whether they expected to go to college when they were in grades 8 to 11.

debates because it reflects the very ability (or inability) of schools to impart skills and knowledge. It is also important because scores on such tests are likely to be highly correlated with scores on the tests routinely used by schools to assign students to particular tracks or programs, and by admissions committees of institutions of higher education. Whatever their intrinsic significance, test scores are then important indirectly: a low score means that the student is more likely to be assigned to a slow track, to find access to higher education more difficult, and so forth, starting a cumulative process leading ultimately to low educational and occupational achievement.

From the available tests, we selected three tests of fundamental skills: mathematics, reading, and vocabulary. Average scores for each group are presented in table 2.3. The scores used are for the subset of items that were identical for sophomores and seniors, so comparisons across cohorts are possible. The scores consist of the total number of items answered correctly, with maxima of 18 for mathematics and 8 for reading and vocabulary. Students who did not attempt to answer any item of a given test are excluded from the tabulation. Addditional analyses, not reported here, of the patterns of nonresponses show that once respondents have begun taking a test it is very likely that they will attempt an answer to every item, and that the small percentages of respondents who do not attempt to give an answer to every item do not differ much across ethnic or racial groups. Therefore, it is unlikely that the patterns of group differences found in table 2.3 are due to different propensities to refuse to take the tests entirely, or to skip a certain number of items.

Table 2.3.--Cognitive achievement scores on mathematics, reading, and vocabulary tests by population subgroup: Spring 1980

Subgroup	Mathematics			Reading			Vocabulary		
	Sample size	Mean score	Standard deviation	Sample size	Mean score	Standard deviation	Sample size	Mean score	Standard deviation
Sophomores									
Mexican-American ..	1,864	7.5	3.5	1,865	2.7	1.7	1,862	2.9	1.6
Cuban	259	8.7	4.3	248	3.5	2.1	254	3.4	2.1
Puerto Rican	313	7.1	3.2	311	2.7	1.8	316	3.0	1.6
Other Latin American	659	8.0	3.4	660	3.0	1.8	659	3.2	1.8
Non-Hispanic black.	868	6.7	3.2	873	2.5	1.7	872	2.7	1.6
Non-Hispanic white.	930	10.3	3.8	931	3.9	2.0	933	4.1	1.9
Seniors									
Mexican-American ..	1,621	8.4	4.0	1,632	3.3	1.9	1,628	3.5	1.8
Cuban	286	10.1	4.3	292	3.9	2.1	292	4.2	1.9
Puerto Rican	257	8.0	4.6	262	3.3	2.0	265	3.5	1.9
Other Latin American	557	8.3	3.9	565	3.3	1.9	567	3.6	1.9
Non-Hispanic black.	854	7.7	3.8	854	3.2	2.0	856	3.2	1.8
Non-Hispanic white.	893	11.6	4.0	901	4.9	2.0	898	4.8	1.9

NOTE: Means and standard deviations are weighted.

For all three tests, whites have the highest scores, blacks the lowest. Among Hispanics, Cubans are ahead of other groups. The scores of Mexican-Americans, Puerto Ricans, and other Latin Americans are similarly low, but higher than the scores of blacks. The high scores of Cubans on the reading and vocabulary tests are especially interesting given the fact, to be documented later, that this subgroup exhibits the highest degree of retention of the Spanish language among Hispanics.

To summarize briefly this review of various indicators of achievement, one could say that clear differences emerge among Hispanic subgroups. The clearest pattern is the tendency for Cubans to perform better than other Hispanics. This is especially true with respect to educational aspirations and the three cognitive achievement scores. In the following sections, we examine several factors that might explain this relative advantage.

2.2. Family Background

In classical analyses of achievement in the tradition initiated by Blau and Duncan (1967), both educational and occupational achievements are viewed as determined in part by socioeconomic characteristics of the family. The factors most often studied in the general literature on this topic, such as occupational prestige, income, and the education of parents, might be called general: they can be measured for all individuals independently of their class, race, or culture. Poverty, for example, cuts across racial and ethnic boundaries. In this section, we examine two of these general background factors: father's education and family income. Other factors, which might be called specific, are

only relevant for some minorities. Examples are the timing of immigration and patterns of language use. We will discuss some of these specific factors later in this report. As we will argue, it is important for policy research to distinguish between the effects of general and specific factors on the fate of Hispanics in school.

2.2.1. Father's Education

Table 2.4 presents the distribution of father's education by population subgroup. As in table 2.2, categories intermediate between high school graduation and college completion have been collapsed for clarity. A salient feature of the table is the fact that the percentage of fathers who did not finish high school is higher for all Hispanic subgroups, although the percentage for other Latin Americans is close to that for whites. These percentages are even considerably higher than those for blacks. An intriguing result is the relatively high percentage for Cubans. This, together with the relatively high proportion of Cuban fathers with college degrees compared with Mexican-Americans and Puerto Ricans, suggests that the educational distribution of the Cuban parental generation is "elongated" compared to that of whites, as if it were composed of two separate components: a group with low educational achievement and a highly educated professional elite. The same may be the case for Other Latin Americans. Puerto Ricans and Mexican-Americans both have very low percentages in all three college categories, consistently below the corresponding figures for blacks.

Note the high percentages of students who do not live with their fathers for both Puerto Ricans and blacks. While for blacks it might

Table 2.4.--Percent distribution of father's education by subgroup: Spring 1980

Subgroup	Sample size	Total	Not residing with father	Less than high school	At least high school, but less than 4 years college	4 year college degree	Master's degree	Ph.D., M.D. or other advanced degree	Total college 1/	Don't know
Sophomores										
Mexican-American	2,009	100.0	10.0	34.5	26.9	2.7	1.4	0.5	4.6	23.9
Cuban	283	100.0	10.5	25.7	35.1	5.2	2.9	2.2	10.3	18.4
Puerto Rican	333	100.0	14.2	27.7	29.7	2.6	1.3	0.4	4.3	24.2
Other Latin American	684	100.0	8.9	14.5	39.1	8.2	3.3	3.1	14.6	23.0
Non-Hispanic black	893	100.0	18.7	16.0	28.5	4.3	2.7	0.8	7.8	29.1
Non-Hispanic white	955	100.0	6.4	13.8	41.0	11.4	6.4	3.3	21.1	17.5
Seniors										
Mexican-American	1,798	100.0	7.6	38.9	31.0	2.8	3.0	1.1	6.9	15.7
Cuban	320	100.0	3.2	36.7	35.4	6.2	4.2	3.6	14.0	10.8
Puerto Rican	278	100.0	18.8	34.9	21.9	1.8	1.7	1.9	5.4	19.0
Other Latin American	607	100.0	9.7	21.6	37.4	8.1	3.3	2.9	14.3	17.1
Non-Hispanic black	909	100.0	20.9	22.0	31.1	3.4	2.0	1.1	6.5	19.6
Non-Hispanic white	968	100.0	6.1	16.7	46.1	12.4	5.8	3.8	22.0	9.1

NOTE: Percentages are weighted.

1/ This column is not included in the percent distribution.

correspond to the high proportion of families headed by females, for Puerto Ricans part of the reason may be the unconventional living arrangements associated with the process of "commuting" migration back and forth between Puerto Rico and the continent. This issue certainly deserves further study.

Finally, note that the percentages in the less-than-high-school category are larger for seniors than for sophomores in all groups. Since fathers of seniors would be expected to be, on the average, roughly two years older than fathers of sophomores, part of this difference may be due to the greater graduation rates of the younger parental cohort, since a large proportion of the students' fathers were of school age at a time of rapid expansion of secondary education in the United States (Duncan, 1965). However, an alternative explanation is that seniors are simply better informed about the educational achievement of their fathers, as the decrease in the Don't-know category from sophomores to seniors for all groups seems to indicate.

2.2.2. Family Income

The distribution of family income (grouped in three categories) is presented in table 2.5. Puerto Ricans appear to be the most disadvantaged subgroup, with 48.6 percent of families having annual incomes less than $12,000. Mexican-Americans are next lowest among Hispanics, although they are somewhat better off than blacks. Predictably, Cubans appear to be the most advantaged, except for whites. There are only 20.4 percent of Cuban families with incomes less than $12,000 and fully 41.5 percent are earning over $20,000, as compared to 10.7 percent and 48.7

Table 2.5.--Percent distribution of yearly family income by population subgroup: Spring 1980

Subgroup	Sample size	Total	Under $12,000	$12,000 to $20,000	Over $20,000
Sophomores					
Mexican American	1,597	100.0	34.9	42.8	22.4
Cuban	252	100.0	25.8	44.7	29.6
Puerto Rican	269	100.0	41.8	44.5	13.7
Other Latin American ..	568	100.0	21.7	44.4	33.9
Non-Hispanic black	714	100.0	36.9	40.4	22.7
Non-Hispanic white	828	100.0	15.4	40.1	44.6
Seniors					
Mexican American	1,598	100.0	29.5	41.5	29.0
Cuban	293	100.0	20.4	38.1	41.5
Puerto Rican	243	100.0	48.6	28.8	22.6
Other Latin American ..	513	100.0	22.5	42.3	35.2
Non-Hispanic black	766	100.0	39.7	35.8	24.5
Non-Hispanic white	871	100.0	10.7	40.5	48.7

NOTE: Percentages are weighted.

percent, respectively, for whites. The Other Latin American income distribution is similar to the one for Cubans. These findings are consistent with current Population Survey results presented in NCES, 1980, p. 28. The results show that, while the proportion of Cubans and Other Latin Americans with incomes below $5,000 is similar to that of the other subgroups, the proportions with incomes above $25,000 is roughly twice that for the other groups.

2.3. Immigration History

Among determinants of educational achievement, the length of contact between an individual and his/her family and U.S. society is

what we have called a specific factor: it has no relevance for populations that have long been settled in the country. As will be shown, Hispanic subgroups differ considerably in immigration history.[1/] These differences may affect achievement in that the acquisition of the skills, values, and lifestyle that characterize U.S. society, and presumably affect the chances of success in that society and its schools, seem to be a function of the length of contact. Length of contact may be defined both intra- and intergenerationally: both the personal experience of the respondent, and the experiences of his/her family may affect achievement.

2.3.1. Nativity

The simplest indicator of contact with U.S. society is the nativity of the respondent: whether he/she was born in the U.S. Table 2.6 presents the distribution of nativity by subgroup. For black and white seniors, mostly long-time resident populations, the proportion born outside the U.S. is very small (4 percent and 2.6 percent, respectively). However, figures vary considerably for Hispanic subgroups: fully 89.4 percent of Mexican-American seniors were born in the U.S. as compared to only 48 percent of Cubans. This difference corresponds to the well-known historical circumstances in which the bulk of Cuban immigration took place. A comparison with the Cuban domestic nativity rate for sophomores, which is 52.6 percent, confirms the pattern: the younger cohort contains more native-born individuals. While absolutely high (79 percent), domestic

---

1/ We use the term "immigrant" throughout this report in the general sense, meaning somebody "who passes or comes into a new habitat or place of residence." (Random House Dictionary of the English Language, Unabridged Edition). The term as used here, therefore, does not refer to the causes of immigration (e.g. economic versus political), or to the legal status of the person in the United States (e.g., "immigrant" versus "refugee").

Table 2.6.--Percent distribution of nativity by population subgroup: Spring 1980

Subgroup	Sample size	Total	Born in U.S.	Not born in U.S.
Sophomores				
Mexican-American	2,077	100.0	88.2	11.8
Cuban	302	100.0	52.7	47.3
Puerto Rican	367	100.0	76.7	23.3
Other Latin American	713	100.0	82.0	18.0
Non-Hispanic black	970	100.0	95.7	4.3
Non-Hispanic white	990	100.0	98.5	1.5
Seniors				
Mexican-American	1,876	100.0	89.4	10.6
Cuban	328	100.0	48.0	52.0
Puerto Rican	299	100.0	79.0	21.0
Other Latin American	629	100.0	82.6	17.4
Non-Hispanic black	972	100.0	96.0	4.0
Non-Hispanic white	991	100.0	97.4	2.6

NOTE: Percentages are weighted.

nativity for Puerto Ricans is lower than that for Mexican-Americans. Note that this figure may be inflated by a possible confusion over the meaning of "United States" in the Puerto Rican context. Other Latin Americans are similar to Puerto Ricans.

2.3.2. Length of Residence of Respondent

Nativity is only a rough indicator of contact with U.S. society since it can conceal substantial differences in immigration histories

of individuals. A student born abroad may have spent almost all of his/her life in the U.S. or only a short period of time. Conversely, a native-born individual may have spent a considerable amount of time outside the continental U.S. This is especially likely to be the case with a pattern of "commuting" migration such as the one that characterizes segments of the Puerto Rican and Mexican communities. A more refined indicator of length of contact that controls for these possibilities is the length of residence in the United States. Table 2.7 presents the relevant figures.

The patterns seen for nativity hold here. An overwhelming majority (89.4 percent) of Mexican-Americans have spent all or almost all of their lives in the United States, while only 55.5 percent of Cubans have done so. For Cubans, the increase from a U.S. nativity rate of 48 percent to a lifetime residency rate ("all or almost all") of 55.5 percent corresponds to the fact that many respondents in this age bracket were brought to this country at an early age during the Cuban influx. For Puerto Ricans, the percentage in the lifetime residency category is about the same (76.3 percent) as that in the domestic-nativity category. It does not follow that all of these respondents are actually native-born: the figures may contain substantial numbers of respondents who were brought to the continental United States at an early age. Native-born Puerto Ricans who have spent a substantial fraction of their lives on the island should be found in the other categories. Note that the largest percentages of very recent immigrants (one to five years) are found for Puerto Ricans (5.3 percent) and Other Latin Americans

Table 2.7.--Percent distribution of length of U.S. residence by population subgroup--how much of the student's life spent in the U.S.: Spring 1980

Subgroup	Sample size	Total	All or almost all	Over 10 years but not all	6 to 10 years	1 to 5 years
Sophomores						
Mexican-American	2,081	100.0	86.9	5.7	3.7	3.7
Cuban	304	100.0	62.1	18.4	17.9	1.7
Puerto Rican ...	367	100.0	75.3	11.5	9.5	3.7
Other Latin American	716	100.0	82.6	6.2	3.5	7.7
Non-Hispanic black	969	100.0	93.2	3.2	2.2	1.5
Non-Hispanic white	992	100.0	97.2	2.1	0.5	0.2
Seniors						
Mexican-American	1,879	100.0	89.7	4.3	3.4	2.6
Cuban	327	100.0	55.5	20.8	21.5	2.2
Puerto Rican ...	299	100.0	76.3	13.8	4.7	5.3
Other Latin American	628	100.0	82.2	6.3	5.9	5.6
Non-Hispanic black	974	100.0	93.7	3.3	1.5	1.5
Non-Hispanic white	991	100.0	97.4	2.2	0.3	0.1

NOTE: Percentages are weighted.

(5.6 percent). This finding for Other Latin Americans, together with the high rate of native birth, suggests that the group is heterogeneous, comprising populations with very different immigration histories.

2.3.3. Length of Residence of Mother

As pointed out above, it is important to take into account intergenerational aspects of length of contact with U.S. society: the experience of a child born in the United States to a family of recent immigrants is

certainly quite different from the experience of a child born to long-established residents. We therefore present, in table 2.8, the distribution for the length of residence of the respondents' mothers. (The patterns for fathers' length of residence were very similar; we present the data for mothers because of their presumed greater import in the development of the linguistic habits of the child.) For Mexican-Americans, note first the percentage in the "all-or-almost-all" category for residency of mother (72.6 percent) compared to the corresponding figure for the length of residency of the respondent (89.7 percent; see table 2.7). This difference means that a large fraction of Mexican-American students, even though native-born or long-time residents, have parents who have not spent all or almost all of their lives in the United States. For Cubans, the distribution again reflects well the historical events, with the mode (39.9 percent) in the 11-to-20-year bracket. Puerto Ricans present a pattern of relatively recent immigration, with only 28.7 percent of mothers in the "all-or-almost-all" category. By contrast, Other Latin Americans, with 72.3 percent, appear to be a relatively long settled population. However, one would expect this group to be quite heterogeneous with respect to immigration history, as discussed in chapter 1.

Length of mother's residence is one variable for which we found substantial differences between the weighted and unweighted figures for Cubans. For example, combining the percentages in the categories 11 to 20 and 6 to 10 years spent in the United States from table 2.8, one obtains a total of 61.6 percent for seniors; performing the same

Table 2.8.--Percent distribution of length of U.S. residence of mother by population subgroup--how much of mother's life spent in U.S.: Spring 1980

Subgroup	Sample size	Total	All or almost all	Over 20 years but not all	11 to 20 years	6 to 10 years	1 to 5 years	Don't know
Sophomores								
Mexican-American ...	2,069	100.0	69.7	10.4	8.4	3.6	1.5	6.5
Cuban	297	100.0	25.5	10.8	41.6	17.6	1.3	3.2
Puerto Rican	365	100.0	37.8	23.1	24.0	3.9	3.5	7.6
Other Latin American.	712	100.0	69.6	7.0	8.9	5.6	3.3	5.5
Non-Hispanic black ..	968	100.0	82.2	5.0	3.0	2.2	0.5	7.0
Non-Hispanic white ..	989	100.0	92.3	3.7	1.4	0.4	0.2	2.0
Seniors								
Mexican-American ...	1,875	100.0	72.6	12.7	6.2	3.4	1.2	3.9
Cuban	325	100.0	12.8	18.4	39.9	21.8	2.3	4.9
Puerto Rican	300	100.0	38.7	30.4	19.5	2.6	4.0	4.7
Other Latin American.	621	100.0	72.3	6.6	10.5	3.8	3.0	3.6
Non-Hispanic black ..	970	100.0	88.5	4.1	4.0	0.4	0.6	2.4
Non-Hispanic white ..	984	100.0	91.2	5.0	2.0	0.4	0.3	1.0

NOTE: Percentages are weighted.

operation with the unweighted figures, the result is 72.6 percent. This seems to indicate that statistical weighting for the Cuban subsample tends to increase the representation of Cubans who have been settled in the country for a long time and who are not part of what most people think of as the Cuban influx of the early sixties. (See the discussion in chapter 1, section 1.4.) The same trend is evident from other comparisons of figures; for example, the proportions of mothers having spent all or almost all of their lives in the United States, for which the unweighted figures are only 5.9 and 13.1 percent for seniors and sophomores, respectively, as compared to the weighted percentages of 12.8 and 25.5. In view of the discussion in chapter 1, this factor should be taken into account when decisions about weighting are made by researchers working with the High School and Beyond sample of Cubans.

2.4. Language

Language skills are obviously an important factor influencing the attainment of culturally distinct minorities. Many policy debates, such as those concerning bilingual-bicultural education, focus on the difficulties of students with limited English-speaking ability in a school system in which English is the dominant language. Language factors are typical of the determinants of achievement specific to certain minorities, although the recent controversy concerning the difficulties faced by students speaking "Black English" might point to a more general problem (see Baratz, 1973). Given the importance of the language factor for Hispanics, we discuss five aspects of language use in this section: home bilingualism, individual bilingualism, Spanish and English proficiency, and frequency of Spanish use.

2.4.1. Home Language

Among the detailed language questions in the High School and Beyond questionnaire was one that asked respondents about both the language usually spoken "by the people at home," and the other language these people speak. On the basis of these two questions it is possible to construct, following Veltman (1979), a typology of home bilingualism with four categories reflecting decreasing degrees of linguistic assimilation of the family into English.[1/] Homes where only English is spoken are considered English monolingual. Homes where English is the usual language but Spanish also is spoken are English-dominant bilingual. Spanish-dominant bilingual homes are those in which Spanish is the usual language, but English is also spoken. Finally, a home is coded as Spanish monolingual when the usual language is Spanish and no other language is spoken. We label this variable home language type.[2/]

1/ Veltman (1979) uses the term "anglicization" to denote the process of linguistic assimilation into English. It is derived from the verb "to anglicize," which is defined by the Random House Dictionary of the English language as "to make or become English in form or character. . . ." Although the term is very useful to denote the linguistic assimilation of language minorities in the United States, we do not use it here because it is not yet standard terminology. We use the term "linguistic assimilation" instead. It should be clear that "assimilation" here refers to language only. We do not discuss other aspects of minority cultures in this report.

2/ The High School and Beyond questions about language were administered separately from the main questionnaire, in the Identification Pages; these also collected detailed information on the respondent's addresses and contacts, which was gathered primarily for use in locating respondents for the followup surveys dictated by the longitudinal structure of the study. The language questions were organized with a filter based on five questions: first language spoken, other language spoken before schooling, usual home language, other language spoken at home, usual language of respondent. Respondents answering English (or equivalent) to these five questions were considered English monolinguals, and were asked not to complete the rest of the language questionnaire. Language data about students who did not pass the filter were not coded and assigned a special type of missing value in the file. These respondents are considered English monolinguals in tables 2.9, 2.10, and 2.11. (See chapter 1 for further discussions of Spanish-language-background respondents.)

The distribution of home language type by Hispanic subgroup is presented in table 2.9. We omit the non-Hispanic population subgroups from the table since the figures would either be negligible for categories other than English monolingual or would correspond to a very heterogeneous array of non-English languages.

The most salient feature of the table is the high proportion of English monolingual families in the Other Latin American group (63.7 percent). The proportions are much smaller for other groups, from 20.6 percent for Cubans to 29.8 percent for Mexican-Americans. Puerto Ricans are similar to Mexican-Americans, with 27.1 percent. The proportions of English-dominant bilingual homes follow the same pattern, except for Other Latin Americans: the largest percentage is for Mexican-Americans (38.2 percent), the lowest for Cubans (9.3 percent), and Puerto Ricans are in between (25.2 percent). The proportions of English monolingual and English-dominant bilingual homes can be viewed as indicators of the linguistic assimilation of a group. Note that the ordering among subgroups corresponds closely to the ordering of nativity and length of residence: by and large, the most recent immigrants are the least assimilated. At the other end of the continuum, the proportions of Spanish-monolingual homes can be viewed as an indicator of the language loyalty of the group. Except for Other Latin Americans, the results are consistent: Cubans have the highest proportion of monolingual homes (26 percent), Mexican-Americans the lowest (11.5 percent), and Puerto Ricans are in between (19 percent). Other Latin Americans, with only 6.9 percent Spanish monolingual households, appear again as a population that is largely assimilated linguistically.

Table 2.9.--Percent distribution of home language type by population subgroup: Spring 1980

Subgroup	Sample size	Total	English monolingual 1/	English-dominant bilingual 2/	Spanish-dominant bilingual 3/	Spanish monolingual 4/
Sophomores						
Mexican-American	2,082	100.0	36.2	31.9	20.7	11.1
Cuban	299	100.0	29.3	9.5	37.2	23.9
Puerto Rican	357	100.0	31.2	14.1	38.8	15.9
Other Latin American .	676	100.0	71.1	15.4	9.3	4.3
Seniors						
Mexican-American	1,867	100.0	29.8	38.2	20.5	11.5
Cuban	322	100.0	20.6	9.3	44.2	26.0
Puerto Rican	305	100.0	27.1	25.2	28.4	19.2
Other Latin American .	605	100.0	63.7	19.7	9.6	6.9

NOTE: Percentages are weighted.

1/ English monolingual: people at home usually speak English, no other language.

2/ English-dominant bilingual: people at home usually speak English, also Spanish.

3/ Spanish-dominant bilingual: people at home usually speak Spanish, also English.

4/ Spanish monolingual: people at home usually speak Spanish, no other language.

2.4.2. Mother Tongue

While current language practices at home are presumably an important determinant of achievement for Hispanics, families also make decisions about the language, or languages that they teach to their children. Independent of current home practices, the language(s) a child learned at an early age may continue to affect achievement at later times. It is also possible that the linguistic practices of the family have changed since the respondent was a young child. To capture early childhood language habits we use two questions from the High School and Beyond instrument. The first one is a mother tongue question: "What was the first language you spoke when you were a child?" The second one might be called a "second mother tongue" question: "What other language did you speak when you were a child--before you started school?" We constructed a variable, labeled mother tongue type, by combining answers to the two questions: respondents whose first language was English and who did not learn any other language before going to school are English monolinguals, those whose first language was English but who also spoke Spanish in childhood are English-dominant bilinguals, and so forth. The principle is the same as for the home language type discussed in section 2.4.1.

The distribution of mother tongue type by Hispanic subgroups is presented in table 2.10. The salient patterns for linguistic assimilation are the same as those for home language type. Except for Other Latin Americans, Mexican-Americans clearly use English most, with 46.3 percent in the English monolingual category. Cubans

Table 2.10.--Percent distribution of mother tongue type by population subgroup: Spring 1980

Subgroup	Sample size	Total	English monolingual 1/	English-dominant bilingual 2/	Spanish-dominant bilingual 3/	Spanish monolingual 4/
Sophomores						
Mexican-American	2,103	100.0	49.5	16.6	21.7	12.3
Cuban	304	100.0	27.5	9.2	31.0	32.3
Puerto Rican	364	100.0	32.1	11.0	35.6	21.2
Other Latin American .	689	100.0	76.7	7.8	7.1	8.4
Seniors						
Mexican-American	1,872	100.0	46.3	18.5	21.3	13.9
Cuban	330	100.0	21.4	9.3	29.4	39.9
Puerto Rican	306	100.0	28.4	14.7	33.1	23.8
Other Latin American .	619	100.0	72.8	8.1	8.6	10.5

NOTE: Percentages are weighted.

1/ English monolingual: English first language spoken, no other language spoken before schooling.

2/ English-dominant bilingual: English first language spoken, Spanish also spoken before schooling.

3/ Spanish-dominant bilingual: Spanish first language spoken, English also spoken before schooling.

4/ Spanish monolingual: Spanish first language spoken, no other language spoken before schooling.

use the most Spanish, with only 21.4 percent in this category. Puerto Ricans are close to Cubans, with 28.4 percent. A very high proportion of Other Latin American families raise their children as English-speaking monolinguals (72.8 percent). An interesting pattern emerges from a comparison of table 2.10 (mother tongue type) with table 2.9 (home language type). For the most assimilated subgroup (Mexican-Americans), the incidence of English monolingualism is considerably higher for mother tongue type than for home usage: 46.3 percent versus 29.8 percent. This may mean that a considerable fraction of English-dominant bilingual families educate their children as English monolinguals. Such a finding has been documented by Veltman (1979), on the basis of the Survey of Income and Education data.[1/]

To explore this matter further, we present in table 2.11 a cross-tabulation of mother tongue type by home language type for all Hispanic seniors and sophomores. The table is arranged to give a picture of the intergenerational process of linguistic assimilation: assuming that current home language type is a good indicator of home language usage at the time a respondent was a small child learning to speak, the table predicts patterns of intergenerational language shifts from Spanish to English during early childhood.[2/]

1/ The same difference in English monolingualism between home language type and mother tongue type holds for Other Latin Americans. The percentage differences are not as large as for Mexicans, however, presumably because the proportion of English-monolingual households for Other Latin Americans is already so high.

2/ The assumption that current language use at home adequately represents the situation when the respondent was a child is certainly unrealistic. Patterns of language use have almost certainly changed for most families over a period of about fifteen years, the general trend being a more frequent use of English. Note that the bias, if any, associated with table 2.11 would be in the direction of underestimating the extent of intergenerational language shift, since the home language use questions, had they been asked fifteen years ago, would presumably have produced larger percentages of Spanish-dominant households. Another possible cause of differences between mother tongue type and home language type is a change in composition of the household between childhood and survey date.

Table 2.11.--Percent distribution of mother tongue type[1/] by home language type[2/]: Spring 1980

Home language type	Sample size	Total	Mother tongue type			
			English monolingual	English-dominant bilingual	Spanish-dominant bilingual	Spanish monolingual
Sophomores						
English monolingual ..	981	100.0	93.2	3.8	2.3	0.6
English-dominant bilingual	926	99.9	42.7	31.0	20.1	6.1
Spanish-dominant bilingual	986	100.0	5.5	10.7	49.3	34.5
Spanish monolingual ...	494	100.0	4.5	13.7	38.8	43.0
Seniors						
English monolingual ..	741	100.0	93.5	2.4	1.8	2.3
English-dominant bilingual	942	100.0	45.1	31.1	19.5	4.3
Spanish-dominant bilingual	901	100.0	2.3	15.0	42.9	39.8
Spanish monolingual ...	485	100.0	4.4	13.3	35.5	46.8

NOTE: Percentages are weighted.

1/ Mother tongue type refers to pre-school language usage. Types are grouped as follows:
English monolingual: English first language spoken, no other language spoken before schooling.
English-dominant bilingual: English first language spoken, Spanish also spoken before schooling.
Spanish-dominant bilingual: Spanish first language spoken, English also spoken before schooling.
Spanish monolingual: Spanish first language spoken, no other language spoken before schooling.

2/ Home language type refers to current language usage at home. Types are grouped as follows:
English monolingual: people at home usually speak English, no other language.
English-dominant bilingual: people at home usually speak English, also Spanish.
Spanish-dominant bilingual: people at home usually speak Spanish, also English.
Spanish monolingual: people at home usually speak Spanish, no other language.

The most salient feature of the table is the indication that among these sophomores and seniors, children in English monolingual homes were raised overwhelmingly as English-monolingual individuals (93.5 percent for seniors). That this figure is not 100 percent is probably due in part to the fact that some families in which Spanish was spoken when the respondents were children have since abandoned the use of the language. It is also striking that almost half of the seniors in English-dominant bilingual households have been raised as English monolinguals. This suggests that the dominance of English as the language spoken in the household entails a major shift toward English monolingualism in the next generation. Further evidence of this phenomenon is obtained by comparing rows 2 and 3 of table 2.11 for seniors: students from Spanish-dominant bilingual households are much more likely to have been raised as Spanish monolinguals or Spanish-dominant bilinguals (39.8 plus 42.9 percent = 82.7 percent), as compared to students raised in English-dominant bilingual households (4.3 plus 19.5 percent = 23.8 percent). The same pattern holds for sophomores. These figures suggest that the major precondition for intergenerational maintenance of the Spanish language for these students is its use as the usual home language. Whether English is also spoken does not seem to have much effect on the shift from Spanish to English dominance, as can be seen by comparing rows 3 and 4 of the table: among respondents raised in Spanish-dominant homes, 17.3 percent (15.0 plus 2.3 percent) were raised as English dominant, a figure that is no larger than the 17.7 percent for respondents raised in Spanish-monolingual homes (13.3 plus 4.4 percent).

Table 2.11 also shows that there is a very small fraction of respondents in Spanish-monolingual households who were raised as English monolinguals (4.4 percent). While such an outcome might seem logically impossible, one has to keep in mind the possibility that a Hispanic child raised exclusively in English was later placed in the custody of relatives who only spoke Spanish. Further analyses of these unusual life histories can be done using the High School and Beyond data.

As a final comment, note that table 2.11 also provides some information concerning the irreversibility of the process of language shift: while 93.5 percent of respondents in English-monolingual homes were raised as English monolingual, only 46.8 percent of those in Spanish monolingual homes were raised as Spanish monolingual (comparison of rows 1 and 4). This result strongly illustrates the intergenerational drift toward English. Again, it should be remembered that home language use patterns have almost certainly changed since the respondent was a child, in the general direction of more frequent use of English. If this is the case, table 2.11 underestimates the strength of the intergenerational process of linguistic assimilation.

2.4.3. Spanish Proficiency

An additional indicator of language loyalty for cultural minorities is proficiency in the use of the non-English language. The High School and Beyond instrument contained four questions to assess proficiency in Spanish. Respondents were asked to rate their ability to understand, speak, read, and write Spanish on a point scale from "very well" to

"not at all." Summing the scores on the four questions yields a composite indicator of proficiency, the distribution for which is presented in table 2.12.[1/]

The table shows that Cubans have the highest mean score (13.5), followed by Puerto Ricans (12.6), Other Latin Americans (11.7), and, finally, Mexican-Americans (11.1). Note the high standard deviation for Other Latin Americans, which suggests again the heterogeneity of this group. By and large, the results for proficiency correspond closely to the pattern in the previous tables: Cubans are the most recent immigrants and the least assimilated linguistically, Mexican-Americans are the oldest residents and the most assimilated linguistically. It should be kept in mind, however, that these tables do not control for possible geographical variations in the process of linguistic shift and maintenance. The study by Veltman (1979), for example, shows that Spanish speakers in Texas (presumably a largely Mexican population) have the highest degree of language retention among Spanish-speaking groups. Further analyses are needed to assess these matters. Another factor to keep in mind is that the measurement of Spanish proficiency employed here is based on self-assessment by the respondent. Students answering the questionnaire may be by temperament more or less optimistic about their abilities to use one language or another. It may also be

---

1/ The decision to measure Spanish proficiency by simply adding the scores for the four questions was based on detailed analyses of answers to these questions in the High School and Beyond pretest. (See also Nielsen 1980, appendix D.) It was found then that the four Spanish proficiency questions loaded about equally on a "Spanish" factor in factor analyses involving both Spanish and English proficiency questions (see the next section for a discussion of English proficiency). Given the structure of the language questions, respondents with English-monolingual backgrounds were not asked to answer the questions concerning language proficiency. For those respondents, proficiency in understanding, speaking, reading, and writing Spanish was imputed to be the minimum value, 4, and English proficiency was imputed to be the maximum, 16.

Table 2.12.--Spanish proficiency--composite scale[1/] of self-assessed ability to understand, speak, read, and write Spanish, by population subgroup: Spring 1980[2/]

Subgroup	Sample size	Mean score	Standard deviation
Sophomores			
Mexican-American	1,625	11.0	3.0
Cuban	272	12.8	2.3
Puerto Rican	279	12.2	2.8
Other Latin American .	346	11.5	3.5
Seniors			
Mexican-American	1,515	11.1	3.1
Cuban	303	13.5	2.2
Puerto Rican	269	12.6	2.9
Other Latin American .	333	11.7	3.7

1/ Measures of composite scale were derived as follows: respondents were asked to rate their ability to understand, speak, read, and write Spanish on a four-point scale ("Very well" to "Not at all"). Summing the scores on the four questions yields our composite indicator of proficiency.

2/ Means and standard deviations are weighted.

the case that some respondents simply have very high verbal abilities, which translate into high proficiency in any language. The positive correlation between Spanish and English proficiency, which is discussed in chapter 3, is consistent with this hypothesis. Whatever the processes involved, measurement of Spanish proficiency in some objective manner could yield quite different results.

2.4.4. English Proficiency

Much debate concerning the fate of Hispanics in school revolves around the adjustment of language-minority pupils to a system based on the English language. One important element in the process is the degree to which such students are proficient in the host language. Respondents were asked to assess their ability to understand, speak, read, and write English, in a format similar to the Spanish questions. We constructed a composite scale of English proficiency by combining the scores on the four questions. The distribution of the English proficiency scores is presented in table 2.13.

Table 2.13.--English proficiency--composite scale[1/] of self-assessed ability to understand, speak, read, and write English, by population subgroup: Spring 1980[2/]

Subgroup	Sample size	Mean score	Standard deviation
Sophomores			
Mexican-American	1,659	14.7	1.9
Cuban	274	15.2	1.3
Puerto Rican	281	14.9	1.7
Other Latin American .	353	15.0	1.8
Seniors			
Mexican-American	1,555	14.8	1.8
Cuban	308	15.0	1.6
Puerto Rican	268	14.7	1.9
Other Latin American .	338	14.6	2.2

1/ Measures of composite scale were derived as follows: respondents were asked to rate their ability to understand, speak, read, and write English on a four-point scale ("Very well" to "Not at all"). Summing the scores on the four questions yields our composite indicator of proficiency.

2/ Means and standard deviations are weighted.

Compared with the distribution for Spanish proficiency (table 2.12), mean English proficiency varies little among subgroups: Mexican-Americans, Puerto Ricans, and other Latin Americans have practically the same scores. Cubans report a slightly higher proficiency than other subgroups. Note that Cubans also appear to have best retained their mother tongue. This suggests that proficiency in two languages does not require a trade-off in which proficiency in one language can be increased only at the expense of the other. It is also worth noting that English proficiency scores are systematically higher than the comparable scores for Spanish proficiency. Obviously, these respondents feel more able to use English than Spanish.

2.4.5. Spanish Usage

Our final linguistic indicator represents an attempt to capture the current language practices of a student in a more refined way. To construct the indicator of Spanish usage, we used four questions concerning the frequency ("how often?") with which a respondent speaks Spanish with the mother and father and the frequency with which each parent speaks Spanish to the respondent. The total usage score is the average of the scores on these four questions. We chose the questions involving parents from a larger set including questions on frequency of Spanish usage with grandparents, siblings, storekeepers, and the like, because previous factor analysis of answers to such questions in the High School and Beyond pretest revealed that the factor loadings of the first component were the largest, and approximately equal, for the four questions (Nielsen, 1980, Appendix D).

The mean scores for the groups are presented in table 2.14. As expected, Cubans have the largest mean usage score (3.2), followed by Puerto Ricans (2.7) and Mexican-Americans (2.1). The scores for Other Latin Americans are very similar to those of the Mexican-Americans. Again, the ordering by timing of immigration is apparent.

Table 2.14.--Composite scale[1/] of four indicators of Spanish use, by population subgroup: Spring 1980[2/]

Subgroup	Sample size	Mean score	Standard deviation
Sophomores			
Mexican-American	1,670	2.1	1.4
Cuban	274	3.1	1.2
Puerto Rican	284	2.7	1.2
Other Latin American .	353	2.0	1.3
Seniors			
Mexican-American	1,557	2.1	1.4
Cuban	309	3.2	1.1
Puerto Rican	264	2.7	1.2
Other Latin American .	333	2.1	1.4

1/ The Spanish use scale is the average of four indicators: the frequency with which (1) respondent speaks Spanish to mother, (2) mother speaks Spanish to respondent, (3) respondent speaks Spanish to father, and (4) father speaks Spanish to respondent.

2/ Means and standard deviations are weighted.

2.5. Conclusions

The tables presented in this chapter yield a composite picture of the differences among Hispanic subgroups. By and large, Cubans have the highest level of achievement and come from families of higher socio-economic status. On the whole, the group is also characterized by recent

immigration and a high degree of retention of the Spanish language, with respect to both frequency of use and proficiency. Mexican-Americans, in contrast, are generally older immigrants, are more assimilated into English, have families of lower socioeconomic status, and achieve less well. Puerto Ricans are intermediate between Mexican-Americans and Cubans in many respects. It is tempting to speculate that the effect of the higher socioeconomic status of Cuban families may be sufficient to compensate for the presumed handicap in school associated with loyalty to a non-English language and recent immigration. If this is the case, group differences in achievement can be explained in terms of general mechanisms of achievement and differences in the compositions of the groups with respect to important determinants. There are, however, alternative explanations of group differences. It may be, for example, that the particular historical circumstances in which Cuban immigration took place, with its climate of general acceptance by the host population, the legal status of the Cubans as political rather than economic migrants, and government policies at the time of Cuban settlement (see Rogg, 1974) affect achievement over and above the compositional differences in background characteristics. If this is the case, the greater achievement of Cubans is not adequately explained by background factors, and there is a residual effect of "Cubanness" that captures the idiosyncracy of the group. To assess this possibility, it is necessary to undertake more refined causal analyses of the effects of background factors on achievement. This is the objective of the third chapter of this report.

CHAPTER 3

A BASELINE MODEL OF ACHIEVEMENT

In this chapter, we present a simple model of the educational achievement process for Hispanics. Our purpose is twofold. First, we wish to investigate the relative importance of the basic input factors we identified above--language use and proficiency, family socio-economic status, length of residence in the United States--as determinants of Hispanic educational attainment. These findings will serve as a basis for comparison for future research investigating the effects of school-level variables on Hispanic scholastic achievement, such as exposure to bilingual or bicultural education. Second, we wish to explore whether the differences in the achievement profiles of Hispanic subgroups which were noted in chapter 2 can be explained by subgroup differences in these basic input factors.

Technically, these questions can be answered by means of regression analysis. This method allows us to assess the relative importance of each background factor once all the others have been controlled. We can also determine whether subgroup differences in achievement remain after the effects of the basic input factors have been taken into account by constructing dummy variables for Hispanic subgroups.

We begin by describing the variables used in the analyses. We then inspect the gross relationships among the independent and dependent variables by considering the correlations in table 3.1. In the regression analyses (tables 3.2 and 3.3) we examine the net effects of the background factors on scholastic achievement. Finally, we use the correlations among the errors of prediction for the student achievement variables to suggest omitted determinants of achievement.

Table 3.1--Correlations, means, and standard deviations for models of scholastic achievement for Hispanics: Spring 1980

	Y_1	Y_2	Y_3	Y_4	Y_5	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	Mean	Standard Deviation
Y_1 = School delay		-.143	-.244	-.191	.189	-.009	-.148	.063	.049	-.175	.062	-.088	-.020	-.039	.617	.740
Y_2 = Educational aspirations	-.128		.414	.326	.354	.236	.085	.106	-.218	.279	-.019	.222	.030	.005	14.950	2.460
Y_3 = Math	-.180	.337		.560	.517	.106	.154	-.017	-.066	.274	.122	.181	-.044	-.009	47.020	9.657
Y_4 = Reading	-.116	.281	.505		.525	.089	.219	-.022	-.055	.198	.041	.147	-.035	-.007	47.505	9.508
Y_5 = Vocabulary	-.130	.300	.417	.468		.133	.185	.016	-.095	.277	.038	.204	-.024	.018	47.814	9.530
X_1 = Spanish proficiency ...	.036	.190	.043	.013	.082		-.223	.773	-.469	-.179	-.157	.258	.127	-.233	10.347	4.144
X_2 = English proficiency ...	-.123	.084	.118	.190	.149	-.236		-.340	.293	.204	.013	.007	-.033	.095	15.088	1.624
X_3 = Spanish use	.092	.071	-.029	-.066	-.023	.768	-.339		-.477	-.303	-.104	.270	.104	-.245	2.009	1.566
X_4 = Length of residence ...	.039	-.144	-.065	-.012	-.066	-.409	.236	-.454		.066	.087	-.422	-.103	-.006	.006	.890
X_5 = SES scale	-.163	.297	.256	.225	.291	-.141	.228	-.265	.053		.092	.186	-.149	.187	-.556	.729
X_6 = Sex (male = 1)	.046	-.048	.097	.029	.022	-.112	-.005	-.051	.050	.051		-.064	-.061	.029	.456	.498
X_7 = Cuban	-.018	.174	.124	.128	.165	.199	.018	.236	-.330	.177	-.047		-.113	-.174	.108	.311
X_8 = Puerto Rican	-.011	.006	-.084	-.043	-.049	.072	.015	.068	-.115	-.070	.001	-.101		-.162	.095	.294
X_9 = Other Latin American ..	-.040	.021	.036	.040	.066	-.234	.126	-.238	.003	.193	.025	-.157	-.171		.200	.400
Mean	.648	14.735	44.961	44.840	45.208	9.632	15.135	1.780	.035	-.470	.442	.085	.099	.211		
Standard deviation	.769	2.654	8.676	8.552	8.914	4.151	1.614	1.544	.852	.722	.497	.279	.299	.408		

NOTE: Seniors (N = 2,350) are above the diagonal, sophomores (N = 2,453) are below the diagonal.

3.1. Variables

We chose as our dependent variables three kinds of policy-relevant achievements: school delay, educational aspirations, and achievement scores on standardized tests. School delay is measured by subtracting the modal age of the population of students for each class (15 for sophomores and 17 for seniors) from the student's self-reported age. This approximation is, of course, subject to inaccuracies since the discrepancy between a student's age and the modal age can be due to factors other than the repetition or skipping of a school grade. Early or late starts in school or the relationship between the student's birthday and the day of survey administration probably generate errors that are likely to be fairly random with respect to anything else in our model. We will consider the consequences of these measurement problems in the regression analyses below.

Students' educational aspirations were measured by the item "As things stand now, how far in school do you think you will get?" The responses were then recoded in years.

We chose standardized tests of three fundamental skills as measures of achievement: reading, vocabulary, and mathematics. To facilitate comparisons across grade, the test scores used here are based on the subset of items common to the two grades. All three tests were standardized across grade to have a mean of 50 and a standard deviation of 10 for the entire High School and Beyond test-taking sample. Students were assigned test scores on each test for which they attempted one or more of the constituent items.

We predict these achievement indicators with a set of variables we conceive to be basic input factors to the schooling process. Family

socioeconomic status is a composite scale identical to the measurement in the 1972 NLS study. It is constructed from father's occupation, father's education, mother's education, family income, and a set of items that ask whether the student's family receives a daily newspaper, owns an encyclopedia or other reference books, a typewriter, an electric dishwasher, two or more cars or trucks, more than 50 books, or a pocket calculator; and whether the student has his or her own room. Each item was standardized within grade to a mean of zero and a standard deviation of one. The mean of the non-missing items was then taken for each case to yield the composite socioeconomic measure.

Language use and proficiency were measured by three variables: Spanish use, Spanish proficiency, English proficiency. All three variables are composite scales constructed from four indicators. Spanish use was derived from the frequency with which the respondent speaks Spanish to his or her mother and father and the frequency with which the parents speak to the respondent in Spanish, a total of four items. Each item had five possible responses ranging from "never" to "always or almost always" and was coded from 0 to 4. Factor analyses of the pretest data showed that these four items formed a single factor with equal weight assigned to each item (Nielsen, 1980, Appendix D); we therefore formed the scale by simply taking the mean of the four items. If a student reported information for a single parent, the scale was generated from the remaining two items. The Spanish use scale, then, has a range of 0 to 4 and is coded positively--that is, higher values indicate a higher frequency of Spanish use. We imputed a score of zero on the Spanish-use scale for any student who did not report information on a language other than English.

As in sections 2.4.3. and 2.4.4. above, Spanish and English proficiency are based on the student's self-assessed ability to understand, speak,

read, and write in each language. There were four response categories ranging from "not at all" to "very well" and coded as integers from 1 to 4. Here, too, exploratory factor analyses of the pretest data were done. This showed the English and Spanish proficiency items formed two separate factors, with each of the indicators--understanding, speaking, reading, and writing--contributing equally to them (Nielsen, 1980, Appendix D). Composite indices were constructed by simply summing the four items corresponding to each language. Note that the coding is positive, ranging from a low of 4 (indicating low proficiency) to a high of 16 (indicating high proficiency). Given this method of scale computation, incomplete information would have made the measures substantively uninterpretable. Because complete linguistic profiles were deemed essential to the model formulation, any case with missing information for one or more of the proficiency indicators was excluded from subsequent analyses. As with the Spanish-use scale, we imputed scores for respondents who did not report language information, the minimum score of 4 on the Spanish-proficiency indicator and the maximum of 16 on the English-proficiency scale.

Family length of residence in the United States is a composite scale derived from the length of residence of the respondent and of the respondent's mother and father. Students were asked how much of their mother's and father's lives had been spent in the United States. Each variable was coded in five categories: (1) about 1 to 5 years, (2) about 6 to 10 years, (3) about 11 to 20 years, (4) more than 20 years, but not all, and (5) all or almost all. Categories 1 through 3 were recoded to the midpoint (3, 8, and 15.5 years, respectively). Categories 4 and 5 presented more of a problem because they implicitly refer to the parent's age, for which we have no measure. The values for these two categories were imputed by using the modal age of mother's

childbearing (25) and adding student's modal age (15 for sophomores, 17 for seniors) and assigning that to the fifth ("all or almost all") category; therefore, the values imputed are 40 for sophomores and 42 for seniors. The midpoint of the fourth category then became defined as 29 years for sophomores and 31 years for seniors. This procedure was essentially repeated for father's length of residence, adding 3 years to account for a typical 3-year difference in age between husbands and wives. Thus, the "all or almost all" category for father's length of residence was recoded to 43 years for sophomores and 45 for seniors. The appropriate midpoint of category 4 was then applied (30.5 for sophomores and 32.5 for seniors).

Students were also asked to report how much of their lives they had spent in the United States. The response categories were: (1) about 1 to 5 years, (2) about 6 to 10 years, (3) more than 10 years, but not all, (4) all or almost all. Since available data included the student's age, all the categories were well-defined and recoded as follows: (1) 3 years, (2) 8 years, (3) (10 + student's age)/2, and (4) student's age.[1/] After each of the three length of residence variables was recoded to the midpoints, each variable was standardized and a simple average was taken of all three indicators. If a student did not report information for either or both parents, the mean was computed on the remaining items. This gives a family-level indicator of the number of years of U.S. residence with equal weight assigned to each item.

In the interest of the proper model specification and to explore further the sex difference in delay rates noted in Tables B.1 and B.2, we include

---

1/ If the student's age was not available, it was imputed for use in the student length-of-residence variable as the age for each class, 15 for sophomores and 17 for seniors.

a dummy variable coded 1 for males and 0 for females in the regressions that follow.

Finally, so as to assess whether the observed differences in achievement between Hispanic subgroups can be explained by our baseline model, we include dummy variables for Cuban, Puerto Rican, and Other Latin American origins. Mexican-Americans were chosen as the comparison (base) group for two reasons, the first statistical, the second substantive: because they are the largest group their exclusion minimizes the standard errors of the dummy coefficients, and all comparisons would have the most linguistically assimilated Hispanic subgroup as their referent.

3.2. Analysis

Table 3.1 shows the correlation, means, and standard deviations for the variables in our model. To insure comparability across the equations predicting the five different dependent variables, we have excluded cases that showed a missing response on any of the fourteen variables. All the analyses to follow, then, have been performed on the same 2,453 sophomores and 2,350 seniors.[1/]

We discuss correlation coefficients briefly here because they constitute convenient summaries of the strength of the bivariate relationships between the variables involved, even though we present better measures of the net effect of these variables on achievement (controlling for

---

1/ The total numbers of Hispanic sophomores and seniors are 3,521 and 3,177, respectively. Preliminary analyses have shown that the substantive results remain unchanged if each dependent variable is analyzed separately, that is, if for any particular dependent variable we do not exclude a case if it is missing on any of the other four dependent variables. We analyze the five dependent variables together here for simplicity of presentation, since a single correlation matrix suffices to describe the interrelations among the variables, and to allow other researchers the flexibility to explore further the causal structures of the dependent variables we analyze here.

other factors) in the regression analyses to follow. We consider such a discusson of correlations to be useful because many of the variables utilized here (such as the linguistic indicators) are not usually available in surveys of this nature, so that few researchers have developed an intuitive feel about the order of magnitude of their relationships with other variables, such as indicators of achievement.

The top lefthand corner of table 3.1 displays the correlations among the dependent variables. Consistent with our expectations, school delay is negatively correlated with the other achievement measures, indicating that high-achieving students are less likely to have been left back.[1/] The largest correlations are among the standardized test scores. The largest relationship is between the mathematics and reading tests, followed by the relationship between the two tests of verbal ability--reading and vocabulary. High-achieving students tend to have high educational aspirations, as demonstrated by the moderate correlations with the standardized tests. As is the case with the other achievement measures, the relationship between aspirations and school delay is slight, albeit in the theoretically expected direction: students left behind have lower aspirations.

When we examine the relationships between the independent and dependent variables, we find in each case that the largest zero-order effect is of family socioeconomic status. Students from higher status families tend to be less delayed, have higher educational aspirations, and do better on all three standardized tests.

The pattern for Spanish use is less clear. The correlations between the test scores and the Spanish-use scale are nil, while the relationships

---

1/ As in chapter 2, we focus on seniors in these analyses, discussing sophomores only when the results are substantially different.

with the measures of delay and educational aspirations are a bit larger. The effects on school delay and educational aspirations, however, are in opposite directions. Students who use Spanish more are more likely to be delayed, but they also tend to have higher educational aspirations. This anomaly disappears when we consider the partial effects in the regression analyses to follow.

Sex differences in achievement are small, as the small magnitude of the correlations indicates. The largest difference is for the mathematics test, where males score better than females. Also worthy of note is the sex-delay correlation: for both grades, males are slightly more likely to be delayed. This corresponds to the small sex differences noted in appendix B.

Lastly, the correlations between the dummy variables and the achievement measures agree with the analysis of Hispanic subgroup differences documented in chapter 2. While Puerto Ricans and Other Latin Americans are not very different from Mexican-Americans on these achievement criteria, Cubans have higher educational aspirations and perform better than Mexican-Americans on all three standardized tests. At least part of this advantage could be due to the fact that Cubans come from families of higher socioeconomic status ($r = .186$). This too is explored in the regression analyses.

3.3. School Delay

Tables 3.2 and 3.3 present regressions of the achievement measures on the independent variables for each class. The metric coefficients denote the level of achievement associated with an increment in each variable, holding constant the level of other explanatory variables. As such, they measure the direct unmediated effects of the background characteristics on achievement. The standardized coefficients also

Table 3.2--Coefficients of models of scholastic achievement for Hispanic sophomores: Spring 1980

	School Delay	Aspirations	Mathematics	Reading	Vocabulary
			Metric Coefficients		
Spanish proficiency	-.006 (-1.02)	.169* (8.88)	.264* (4.11)	.253* (3.97)	.434* (6.68)
English proficiency	-.045* (-4.36)	.113 (3.38)	.505* (4.47)	.823* (7.33)	.614* (5.37)
Spanish use	.048* (2.82)	-.144* (-2.63)	-.428* (-2.31)	-.507* (-2.76)	-.621* (-3.32)
Length of residence	.096* (4.42)	-.256* (-3.64)	-.837* (-3.52)	-.266 (-1.13)	-.477* (-1.99)
SES scale	-.146* (-6.29)	1.063* (14.10)	2.658* (10.43)	1.952* (7.71)	2.962* (11.49)
Sex (male = 1)	.077* (2.48)	-.165 (-1.65)	1.804* (5.33)	.623 (1.85)	.596 (1.74)
Cuban	.084 (1.35)	.606* (2.99)	1.335 (1.95)	2.546* (3.74)	3.048* (4.40)
Puerto Rican	-.015 (-.28)	.096 (.56)	-2.368* (-4.08)	-.895 (-1.55)	-.904 (-1.54)
Other Latin American	.030 (.73)	.076 (.57)	-.347 (-.77)	.046 (.10)	.801 (1.77)
Intercept	1.184	12.160	36.212	31.361	33.654
R square	.047	.160	.102	.088	.129

NOTE: T-values are in parentheses; coefficients with asterisks are significantly different from zero at .05 probability level (two-tailed test).

Table 3.2--Coefficients of models of scholastic achievement for Hispanic sophomores: Spring 1980--(Continued)

	School Delay	Aspirations	Mathematics	Reading	Vocabulary
	Standardized Coefficients				
Spanish proficiency	-.032	.264	.126	.123	.202
English proficiency	-.094	.069	.094	.155	.111
Spanish use	.096	-.084	-.076	-.092	-.108
Length of residence	.106	-.082	-.082	-.026	-.046
SES scale	-.137	.289	.221	.165	.240
Sex (male = 1)	.050	-.031	.103	.036	.033
Cuban	.030	.064	.043	.083	.095
Puerto Rican	-.006	.011	-.082	-.031	-.030
Other Latin American	.016	.012	-.016	.002	.037

Table 3.3.--Coefficients of models of scholastic achievement for Hispanic seniors: Spring 1980

	School Delay	Aspirations	Mathematics	Reading	Vocabulary
			Metric Coefficients		
Spanish proficiency	-.012*	.170*	.593*	.475*	.539*
	(-2.11)	(9.44)	(8.15)	(12.42)	(7.54)
English proficiency	-.060*	.159*	.768*	1.277*	1.025*
	(-5.90)	(5.12)	(6.11)	(10.12)	(8.31)
Spanish use	.040*	-.125*	-.840*	-.669*	-.478*
	(2.43)	(-2.48)	(-4.12)	(-3.27)	(-2.39)
Length of residence	.068*	-.391*	-.407	-.691*	-.533*
	(3.12)	(-5.82)	(-1.50)	(-2.54)	(-2.00)
SES scale	-.147*	.967*	3.011*	1.909*	3.106*
	(-6.43)	(13.69)	(10.58)	(6.67)	(11.10)
Sex (male = 1)	.095*	.041	2.605*	1.049*	1.017*
	(3.14)	(.44)	(6.92)	(2.77)	(2.75)
Cuban	-.076	.487*	2.965*	1.921*	3.201*
	(-1.30)	(2.72)	(4.10)	(2.64)	(4.50)
Puerto Rican	-.097	.345*	-.440	-.696	.211
	(-1.83)	(2.11)	(-.67)	(-1.05)	(.33)
Other Latin American	-.015	.032	-.679	-.680	.228
	(-.36)	(.26)	(-1.34)	(-1.33)	(.46)
Intercept	1.459	11.461	31.335	25.247	28.580
R square	.063	.193	.148	.114	.156

NOTE: T-values are in parentheses; coefficients with asterisks are significantly different from zero at .05 probability level (two-tailed test).

Table 3.3.--Coefficients of models of scholastic achievement for Hispanic seniors: Spring 1980 (Continued)

	School Delay	Aspirations	Mathematics	Reading	Vocabulary
			Standardized Coefficients		
Spanish proficiency	-.069	.287	.254	.207	.234
English proficiency	-.131	.105	.129	.218	.175
Spanish use	.084	-.080	-.136	-.110	-.078
Length of residence	.082	-.142	-.037	-.065	-.050
SES scale	-.145	.286	.227	.146	.238
Sex (male = 1)	.064	.008	.134	.055	.053
Cuban	-.032	.062	.096	.063	.104
Puerto Rican	-.038	.041	-.013	-.021	.006
Other Latin American	-.008	.005	-.028	-.029	.010

measure the effects of background factors but in addition have the desirable property of expressing the relationships between the achievement and input variables in standard deviation units. This allows us to directly compare the relative importance of the input factors in determining each achievement variable. Lastly, the coefficients of the dummy variables for Hispanic subgroup can be inspected to test the substantive issue raised in the previous chapter: whether subgroup differences in achievement remain after the effects of the basic input factors have been controlled. Insignificant dummy coefficients would indicate that subgroup differences in background variables account for the differences in achievement across subgroup.

Examining the equation for school delay, we find that, except for the statistically insignificant coefficients of the dummy variables for Hispanic subgroups, the models for sophomores and seniors are in agreement with regard to the signs of the slopes. (The R-square values range from .05 for sophomores to .06 for seniors.)

With everything else in the model controlled, seniors who report themselves as highly proficient in Spanish tend to be less delayed. (This relationship fails to be significant for sophomores.) One possible explanation of this result is that, holding everything else constant, Spanish proficiency is in part an indicator of general verbal ability, which should be positively related to scholastic achievement and, therefore, negatively related to school delay. This interpretation is given some support by the fact that students with high Spanish proficiency score higher on the vocabulary test (see the last column of table 3.3). As expected, there is a somewhat greater tendency for students who are highly proficient in English to be less delayed in school (compare the standardized coefficients). The effect of Spanish use, however, is

in the opposite direction. The more a student uses Spanish, the more likely he or she is to be held behind. Since English proficiency has been controlled in these models, we cannot explain this result simply by the linguistic handicap that Hispanic students might suffer in an English-language educational system.

One possible explanation might be that the Spanish-use scale is in part measuring the deleterious effects of code switching (English to Spanish and vice versa) on achievement. Bilingual respondents might suffer the cognitive costs of maintaining two languages independent of the level of proficiency in either language. The fact that degree of Spanish use is associated with greater school delay even after the levels of both English and Spanish proficiency have been controlled is consistent with this interpretation.

A second interpretation might focus on the institutional context within which Spanish is used. It might be that Hispanic students in school settings that encourage the use of and facility in Spanish (e.g., bilingual/bicultural programs) are less alienated from school and therefore achieve better. However, since relatively few students are to be found in these settings, greater Spanish use appears as negatively related to achievement independent of the levels of Spanish and English proficiency. We will explore this hypothesis in subsequent reports when we consider the effects of bilingual/bicultural programs.

Considering family length of United States residence, we note a highly significant tendency for students whose families have resided in the United States longer to be delayed in their progress through school. This pattern of longer United States residence being associated with lower achievement is evident for the other measures as well, as

the coefficient for length of residence in the other equations demonstrates.

These findings might suggest the presence of a selection process associated with immigration. Since families with higher status are more likely to be able to marshal the resources necessary for immigration, they also bring with them a constellation of attributes that would tend to encourage high academic achievement. The immigration experience of Cubans might serve as a case in point; recall the relatively high status of Cubans as compared with other Hispanic subgroups that we documented in chapter 2. This hypothesis falls short, however, since the effects of family socioeconomic status and being Cuban have been controlled in these models.

Another explanation of these results might be that Hispanic families settled in the U.S. for a long period of time have become more "ghettoized," acquiring the propensity to discouragement and low achievement associated with their marginal status (Kimball, 1968; Baral, 1979). Note, however, that the most likely effect of such marginalization would be dropping out of school altogether. Since our sample is composed of "survivors"--students who have made it to at least the tenth grade (see section 1.4 in chapter 1)--it is likely that the models we present here underestimate the importance of length of residence as a determinant of achievement for the population of school-age Hispanics. In the absence of data for dropouts, we cannot directly test this inference. Clearly, further analyses are needed to uncover the mechanism by which length of residence affects student achievement.

Turning to family socioeconomic status, we note that Hispanic students tend to come from less affluent families, as the means for the SES scale in table 3.1 show. For each class, Hispanics tend to

be half a standard deviation below the general population on the SES scale. Hispanics are also relatively more concentrated around the mean, as the standard deviations indicate.

Despite the lower status of Hispanics relative to the general population, family socioeconomic status surfaces as the most important determinant of school delay for both classes (note the standardized coefficients). With everything else in the model controlled, students from families of higher socioeconomic status are significantly less likely to be delayed in school.

We interpret this result as indicative of two factors that would decrease the likelihood of student delay. First, family socioeconomic status is an indicator of the material resources that the family can devote to advance the child's education. Second, as more affluent families are likely to have highly educated parents, the family SES indicator may be capturing variation in parental attitudes toward education. Since parental education was used in constructing the SES indicator we cannot test this interpretation in the models we present here. Preliminary analyses, however, have shown that family educational milieu as measured by father's education, while fairly highly correlated with family income (approximately .40), is a significant predictor of all five measures of achievement. We have chosen to use the composite SES indicator rather than using separate family income and parental education measures because of the large number of missing values for those variables.

For both grades, we find a sex difference in school delay: males are significantly more likely to be delayed than females. This corresponds to the small differences in delay rates we noted in section 2.1.1 of chapter 2. The fact that this sex difference still appears in the regressions

implies that the causes of the greater propensity of males to be delayed are out of the purview of our baseline model. Whether sex differences in socialization and behavior or school- and classroom-level student evaluation processes will ultimately account for the differences in delay rates we document here is unclear. Subsequent reports will explore these hypotheses.

Finally, the coefficients for the Hispanic subgroup dummy variables are statistically insignificant, indicating that the subgroup differences in delay we found in table 2.1 are explained by other variables in our baseline model. Specifically, the somewhat lower likelihood of delay for Cubans is explained by their greater Spanish proficiency, more recent immigration, and higher socioeconomic status. In the same way, the lower delay rates of Other Latin Americans is explained by their relatively higher socioeconomic background (r = .187).

Inspection of the R-squares shows that, in absolute terms, our linear models are not very effective in explaining the variance in school delay. Previous discussion of variable limitations, however, indicated that the dependent variable in these models is likely to include random measurement error, which, by definition, is not predictable above chance levels. Although this may decrease the explanatory power of our model (the R-squares), it should not bias our unstandardized estimates of the effects of the background factors on school delay.[1/] For example, being

---

1/ Though random measurement error in the dependent variable does not affect the unstandardized regression coefficients, it does attenuate the standardized weights. Even here, however, random measurement error in the dependent variable does not change the relative magnitude of the standardized effect, only the absolute sizes. At worst, then, random error in our measure of delay would force us to underestimate the population values of the standardized coefficients. Since we are at this point concerned with making causal inferences about the factors that affect school delay rather than population inferences, we consider the attenuated R-squares and standarized coefficients a small price to pay.

male is associated with about one-tenth of a year of school delay. Going from the "not very well" to the "pretty well" category on a single item of the English proficiency scale is worth one-twentieth of a year of school advancement. The same one-category shift in Spanish use yields about one-twentieth of a year of school delay for both classes. All these inferences are, of course, based on holding all other variables constant.

The small sizes of these effects is strong evidence that the determinants of school delay for Hispanics are not among the basic input factors we analyze here. This is consistent with the hypothesis that school- and system-level discretionary policies and practices are the major causes of school delay. In subsequent reports, we will explore directly whether variations in school policies explain school delay.

3.4. Aspirations

Turning to the models predicting educational aspirations, we see that the R-squares are considerably higher for aspirations than for school delay. In large part, however, the models for aspirations show the same determinants as those we found for school delay.

High Spanish proficiency is associated with high educational aspirations just as it was related to delay. Inspection of the standardized coefficients shows this to be among the most important determinants of aspirations. Here, too, we interpret this result as being due to Spanish proficiency's role as an indicator of general verbal ability.

The other two language indicators behave as hypothesized. With everything else held constant, students with a better command of English have higher aspirations, while more frequent use of Spanish in communicating with parents is related to lower educational aspirations. The verbal

ability component of the English proficiency indicator would explain the first result. As with school delay, either code switching or the effect of institutional factors might explain the negative relationship between Spanish use and aspirations.

When we consider the effect of family length of residence in the United States, we find that students whose families have been in the U.S. longer have lower aspirations than recent immigrants. One might be tempted to explain this result by imputing immigrants' visions of America as the land of opportunity (but see Portes et al., 1978). These recent immigrants, however, differ not only in aspirations, but also in actual test performance, as demonstrated by the coefficients for length of residence in other equations. As we argued above with regard to school delay, selection processes that occur during immigration are unlikely to account for this result. Here, as for the other achievement measures, we interpret the tendency of more recent immigrants to perform better on achievement criteria as offering tentative support for the ghettoization/marginalization hypothesis.

As expected, students from families with high socioeconomic status have higher aspirations. We hypothesize that the effects of both material affluence and family educational milieu are being captured here. We also note, for both classes, that the sex differences in educational aspirations are insignificant.

While this lack of difference has been found repeatedly in the general population (e.g., Rosen and Aneshensel, 1978; Erbring, 1981), it may be somewhat surprising to those familiar with groups, such as ASPIRA, that are concerned with increasing the proportion of Hispanics in higher education. One complaint that is often heard in casual conversations

concerns the greater difficulty experienced by Hispanic women in going to college, which often involves moving out of the parental home, because of family opposition based on traditional assumptions about proper roles for young women. These results seem to indicate this is not the case: the educational aspirations of Hispanic men and women do not differ significantly. However, the dependent variable here is educational aspirations, not actual achievement. Women at this stage may not yet have experienced family opposition based on traditional values.

Finally, the slopes for the dummy variables for Hispanic subgroups are not consistent across high school grades. For both classes, Cubans are found to have higher aspirations than all other subgroups, even after the effects of the remaining variables are removed, and Other Latin Americans are similar to Mexican-Americans. But Puerto Rican seniors have higher aspirations than Mexican-Americans, though this is not true of Puerto Rican sophomores. We have no explanation for this discrepancy, though we might consider the potential influence of normal passage through high school. Since seniors are closer to an important life transition--into either college or the labor force--their educational aspirations are likely to be more concretely formulated. The presence of more crystallized future ambitions among seniors might allow demographic differences among the subgroups to surface in predicting educational aspirations. The relatively higher R-square for seniors supplies some indirect support for this interpretation.

3.5. Test Scores

Students who are highly proficient in Spanish perform better on all three achievement tests. The same is true of students highly proficient in English. As is the pattern for school delay and educational

aspirations, a high rate of Spanish usage is associated with lower achievement on all three tests.

Among sophomores, more recent immigrants perform significantly better on the mathematics and vocabulary tests. Among seniors, recency of immigration is related to higher achievement on the reading and vocabulary tests. Length of residence is not a significant predictor of performance on the reading test for sophomores and on the mathematics test for seniors. Especially in light of our discussion above regarding the ghettoization hypothesis, we lack a substantive explanation of why these effects do not emerge at conventional levels of statistical significance.

For both classes, the slopes for the SES scale indicate that students of more affluent families tend to perform better on all three tests of basic skills. As noted above, preliminary analyses (not reported here) supported the interpretation of SES as indicative of two conceptually distinct factors that would raise student achievement: (1) higher affluence, which measures the resources that may be devoted to schooling, and (2) parental education, which might serve to engender more favorable attitudes toward school.

Among the subgroups, Cubans score somewhat higher than all others on the mathematics test for sophomores. Puerto Rican sophomores are low achievers on the math test. The remaining slopes for the subgroups are statistically insignificant.

With respect to sex differences in test scores, males perform better than females in both classes, although this tendency is not statistically significant for sophomores on the reading and vocabulary tests.

Finally, consideration of the R-squares reveals that we have been able to explain approximately 10 to 15 percent of the variance in test scores. We might here speculate about the omitted factors that

might contribute to the unexplained variance. Table 3.4 presents the correlations among the residuals for our five achievement variables, which are best interpreted as the relationships between the dependent variables once the effects of all the independent variables have been controlled. As such, they represent a good indicator of how much systematic variance is due to factors not included in the baseline model.

3.6. Residual Correlations

Parallel to the zero-order correlations (see table 3.1), the strongest relationships in table 3.4 are found among the test scores. Compared with the top lefthand corner of table 3.1, the magnitude of the correlations has been reduced, implying that portions of the associations have been explained with our basic input factors. Of course, the absolute sizes of the figures indicate substantial effects of variables not included in the models.

Two mechanisms come to mind as potential explanations of the high correlations between test score errors. First, the unmeasured effects of school-level variables (such as per capita expenditures, schoolwide racial/ethnic mix, and teacher characteristics) might serve as one source of common error. Second, mental capacity or intelligence is one unexplored input factor that might explain part of the relationship between test scores. To the extent that intelligence is uncorrelated with the independent variables subsumed in our basic model, we would expect its independent effect to reduce the correlations among the errors of prediction for the test scores.

In the case of school delay, the smaller R-squares suggest that the background factors we selected are relatively ineffective in predicting this form of achievement. The smaller correlations between the school

Table 3.4--Correlations among the errors for the models of scholastic achievement: Spring 1980

	School Delay	Aspirations	Mathematics	Reading	Vocabulary
School delay		-.065	-.194	-.133	-.119
Aspirations	-.076		.324	.240	.234
Mathematics	-.143	.266		.505	.437
Reading	-.071	.213	.462		.459
Vocabulary	-.078	.195	.350	.412	

NOTE: Seniors above and sophomores below the diagonal.

delay and the other equation errors also suggest that relevant factors omitted from the school delay equation would be likely to have little causal impact on the other forms of achievement. The random error suspected in our measure of delay would behave precisely in this manner. By the very nature of randomness, the error in delay should be uncorrelated with everything, including the variables omitted from the other equations. However, even if we were to prove the existence of random error in our measure of delay, thereby excusing the small R-squares and standardized coefficients, we could not explain the small size of the metric coefficients in our model. This is strong evidence that the determinants of school delay for Hispanics are not among the basic input factors we have analyzed here.

The errors in prediction of educational aspirations are moderately correlated with the errors associated with the test score equations. Interpretations might follow three lines of reasoning.

First, it might be argued that there exists some set of factors, such as intelligence or school-level attributes, that has been omitted from our model and that determines both aspirations and test scores. With only the baseline model before us, we cannot reject this hypothesis.

Second, aspirations might be best considered as a consequence of other forms of achievement. The correlations between the errors in educational aspirations and test-score performance would, therefore, be due in part to the causal impact of achievement on aspirations. Controlling for test score achievement in the aspiration model would leave only that portion of the variance in aspirations that is due to omitted factors. If these same factors are relevant to the equations predicting standardized tests, then the errors in these equations would remain positively correlated.

Third, a more dynamic model might take the relationship one step further and argue that aspirations and achievement are indeed separate concepts, but that they cause each other. In this case, the correlated errors would reflect both causal processes. It will not be possible to resolve the issue until the data from the next wave of High School and Beyond become available.

3.7. Summary and Conclusion

The multivariate analyses contained in this chapter reveal a number of consistent patterns concerning determinants of Hispanic scholastic achievement. Hispanic students from more affluent families performed better on all five achievement measures; i.e. they are less delayed in their progress through school, have higher educational aspirations, and scored higher on the mathematics, reading, and vocabulary tests.

With regard to linguistic factors, those students who are highly proficient in English are better achievers. Interestingly, the same relationship surfaces for Spanish proficiency: students with greater facility in the Spanish language tend to be higher achievers, even after the other variables in the model are controlled. However, those students who use Spanish more frequently appear as lower achievers.

Family length of residence is negatively associated with achievement; i.e., students in families of more recent immigration achieve higher than those in families of long-time residence. Gender differences in achievement for these Hispanic 10th and 12th graders are inconsistent. After controlling the other input factors, males tend to be more delayed than females, yet, at least among seniors, males tend to perform better than females on the mathematics, reading, and vocabulary tests.

Lastly, the causal analyses of various aspects of the achievement of Hispanics presented above allow some insights into the structure of group differences found in chapter 2 of this report. In particular, the coefficients of the dummy variables corresponding to Hispanic subgroups provide measures of the extent to which group differences in achievement persist after we have controlled for a variety of causal factors. Generally speaking, differences between Puerto Ricans and Mexican-Americans seem well explained by the background factors alone: the coefficient of the Puerto Rican dummy variable is significant only in the equation for the aspirations of seniors and the equation for the mathematics score of sophomores. This indicates that the mere fact of being Puerto Rican or Mexican-American explains relatively little variation in achievement over and above background characteristics. In contrast, the coefficients of the Cuban dummy variable are significant in seven out of ten equations. In other words, "Cubanness" enhances achievement relative to Mexican-Americans (the reference category) independently of the set of explanatory factors that we have selected. On the basis of these results, one could argue that the particular situation of Cubans as a group in the United States, such as the climate of general political sympathy at the time of their settlement, favors the educational achievement of these students. However, the issue is far from settled in view of the deliberate simplicity of our baseline model: it is quite possible that introducing additional independent variables would cause the "Cuban difference" to vanish.

BLANK PAGE
DO NOT FILM

APPENDIX A

SELECTED QUESTIONS ON LANGUAGE USE, LANGUAGE PROFICIENCY, AND NATIONAL ORIGIN

NOTE: The language questions on the following pages are taken from the Student Identification Pages. These questions were identical for the two cohorts. The national origin question appears as question 90 in the senior questionnaire and question 91 in the sophomore questionnaire.

BLANK PAGE
DO NOT FILM

The following questions are about the language or languages spoken by you and your family.

11. What was the first language you spoke when you were a child? (MARK ONE)

- English ..... ○ 01
- Spanish ..... ○ 02
- Italian ..... ○ 03
- Chinese ..... ○ 04
- French ..... ○ 05
- German ..... ○ 06
- Greek ..... ○ 07
- Portuguese ..... ○ 08
- Filipino languages ..... ○ 09
- Polish ..... ○ 10

Other: (Write in) _____ ○ 11

---

12. What other language did you speak when you were a child—before you started school? (MARK ONE. IF MORE THAN ONE, MARK ONE MOST OFTEN SPOKEN.)

I spoke no other language ..... ○ 01

I also spoke:

- English ..... ○ 02
- Spanish ..... ○ 03
- Italian ..... ○ 04
- Chinese ..... ○ 05
- French ..... ○ 06
- German ..... ○ 07
- Greek ..... ○ 08
- Portuguese ..... ○ 09
- Filipino languages ..... ○ 10
- Polish ..... ○ 11

Other: (Write in) _____ ○ 12

---

13. What language do you usually speak now? (MARK ONE)

- English ..... ○ 01
- Spanish ..... ○ 02
- Italian ..... ○ 03
- Chinese ..... ○ 04
- French ..... ○ 05
- German ..... ○ 06
- Greek ..... ○ 07
- Portuguese ..... ○ 08
- Filipino languages ..... ○ 09
- Polish ..... ○ 10

Other: (Write in) _____ ○ 11

---

14. What language do the people in your home usually speak? (MARK ONE)

- English ..... ○ 01
- Spanish ..... ○ 02
- Italian ..... ○ 03
- Chinese ..... ○ 04
- French ..... ○ 05
- German ..... ○ 06
- Greek ..... ○ 07
- Portuguese ..... ○ 08
- Filipino languages ..... ○ 09
- Polish ..... ○ 10

Other: (Write in) _____ ○ 11

---

15. What other language is spoken in your home? (MARK ONE. IF MORE THAN ONE OTHER LANGUAGE IS SPOKEN, MARK THE OTHER LANGUAGE WHICH IS SPOKEN MOST OFTEN.)

No other language is spoken ..... ○ 01

The other language spoken is:

- English ..... ○ 02
- Spanish ..... ○ 03
- Italian ..... ○ 04
- Chinese ..... ○ 05
- French ..... ○ 06
- German ..... ○ 07
- Greek ..... ○ 08
- Portuguese ..... ○ 09
- Filipino languages ..... ○ 10
- Polish ..... ○ 11

Other: (Write in) _____ ○ 12

---

16. Please look back at your answers to Questions 11 - 15 . . .

. . . IF you answered ENGLISH (or no other language) to ALL FIVE QUESTIONS, you have completed this section of the questionnaire. Thank you.

. . . IF you answered a LANGUAGE OTHER THAN ENGLISH IN ANY OF THE FIVE QUESTIONS,* please write the name of that language here _____ then CONTINUE with the rest of this questionnaire. Most of the questions that follow are about the use of that language by you and your family.

*IF YOU ANSWERED MORE THAN ONE NON-ENGLISH LANGUAGE in Questions 11 - 15 please write the most important one on the line.

17. With regard to that language, how well do you do the following? (MARK ONE OVAL FOR EACH LINE)

How well do you . . .	Very Well	Pretty Well	Not Very Well	Not at All
a. Understand that language when people speak it	O	O	O	O
b. Speak that language	O	O	O	O
c. Read that language	O	O	O	O
d. Write that language	O	O	O	O
	1	2	3	4

18. How often is that language spoken by the person underlined in each of the situations listed below? (MARK ONE OVAL FOR EACH LINE. IF YOU DO NOT LIVE WITH THE RELATIVE INDICATED OR DO NOT SEE THAT PERSON OFTEN, PLEASE MARK THE OVAL UNDER "Does not apply.")

How often do (does):	Always or almost always	Mostly	About half the time	Sometimes	Never	Does not apply
a. You speak that language to your mother	O	O	O	O	O	O
b. Your mother speak that language to you	O	O	O	O	O	O
c. You speak that language to your father	O	O	O	O	O	O
d. Your father speak that language to you	O	O	O	O	O	O
e. Your parents speak that language to each other	O	O	O	O	O	O
f. Other relatives (brothers sisters, grandparents) speak that language around you	O	O	O	O	O	O
g. You speak that language with your best friends	O	O	O	O	O	O
h. You speak that language in school with other students	O	O	O	O	O	O
i. You speak that language in the stores you go to most often (i.e., grocery, record store, clothes store)	O	O	O	O	O	O
j. You speak that language at work	O	O	O	O	O	O
	01	02	03	04	05	06

19. How well do you do the following? (MARK ONE OVAL FOR EACH LINE)

	Very Well	Pretty Well	Not Very Well	Not at All
a. Understand spoken English	O	O	O	O
b. Speak English	O	O	O	O
c. Read English	O	O	O	O
d. Write English	O	O	O	O
	1	2	3	4

EDUCATION IN THE UNITED STATES

This series of questions concerns subjects you may have had in school. Please answer only for education you have received in the United States.

20. Did you have the following courses in grades 1 - 6? (MARK ONE OVAL FOR EACH LINE)

Did you have . . .	Yes	No	Not in U.S. in grades 1 - 6
a. An English course designed for students from non-English speaking backgrounds	O	O	O
b. Reading and writing in that language (refer to Q. 16 for "that language")	O	O	O
c. Other subjects, such as math or science, taught, at least in part, in that language	O	O	O
d. Courses in the history and culture of your ancestors' country of origin or their life in the United States	O	O	O
	1	2	3

21. Did you have the following courses in grades 7 - 9? (MARK ONE OVAL FOR EACH LINE)

Did you have . . .	Yes	No	Not in U.S. in grades 7 - 9
a. An English course designed for students from non-English speaking backgrounds	O	O	O
b. Reading and writing in that language (refer to Q. 16 for "that language")	O	O	O
c. Other subjects, such as math or science, taught, at least in part, in that language	O	O	O
d. Courses in the history and culture of your ancestors' country of origin or their life in the United States	O	O	O
	1	2	3

22. Did you have the following courses in grades 10 - 12? (MARK ONE OVAL FOR EACH LINE)

Did you have . . .	Yes	No
a. An English course designed for students from non-English speaking backgrounds	O	O
b. Reading and writing in that language (refer to Q. 16 for "that language")	O	O
c. Other subjects, such as math or science, taught, at least in part, in that language	O	O
d. Courses in the history and culture of your ancestors' country of origin or their life in the United States	O	O
	1	2

---

23. Thinking about all the courses you had in each of those grades listed below, how much of the teaching was done in that language?

A. In grades 1 - 6: (MARK ONE)

All or almost all of the teaching was done in that language	O 01
Most was in that language	O 02
About half was in that language	O 03
Some was in that language	O 04
None was in that language	O 05
Was not in school in U.S. then	O 06

B. In grades 7 - 9: (MARK ONE)

All or almost all of the teaching was done in that language	O 01
Most was in that language	O 02
About half was in that language	O 03
Some was in that language	O 04
None was in that language	O 05
Was not in school in U.S. then	O 06

C. In grades 10 - 12: (MARK ONE)

All or almost all of the teaching was done in that language	O 01
Most was in that language	O 02
About half was in that language	O 03
Some was in that language	O 04
None was in that language	O 05

---

91. What is your origin or descent? (If more than one, please mark below the one you consider the most important part of your background.) (MARK ONE)

HISPANIC OR SPANISH:

- Mexican, Mexican-American, Chicano ..... O
- Cuban, Cubano ..... O
- Puerto Rican, Puertorriqueno or Boricua ..... O
- Other Latin American, Latino, Hispanic, or Spanish descent ..... O

NON-HISPANIC:

- African:
  - Afro-American ..... O
  - West Indian or Carribean ..... O
- Alaskan Native ..... O
- American Indian ..... O
- Asian or Pacific Islander:
  - Chinese ..... O
  - Filipino ..... O
  - Indian, Pakistani or other South Asian ..... O
  - Japanese ..... O
  - Korean ..... O
  - Vietnamese ..... O
  - Other Pacific Islander ..... O
  - Other Asian ..... O
- European:
  - English or Welsh ..... O
  - French ..... O
  - German ..... O
  - Greek ..... O
  - Irish ..... O
  - Italian ..... O
  - Polish ..... O
  - Portuguese ..... O
  - Russian ..... O
  - Scottish ..... O
  - Other European ..... O
- Canadian (French) ..... O
- Canadian (Other) ..... O
- United States only ..... O

Other (WRITE IN) _____ ..... O

APPENDIX B

SUPPLEMENTARY TABLES: AGE DISTRIBUTIONS BY SEX

Table B.1.--Percent distribution of age by population subgroup for males: Spring 1980

Subgroup	Sample size	13 or younger	14	15	16	17	18	19	20	21 and older	At least 2 years delayed 1/
Sophomores											
Mexican-American	899	0.4	0.5	38.7	44.9	12.8	1.8	0.6	--	0.2	15.5
Cuban	123	--	2.3	37.5	44.8	12.8	--	2.6	--	--	15.4
Puerto Rican	151	--	2.6	39.4	42.3	14.2	1.5	--	--	--	15.7
Other Latin American	308	1.1	0.5	42.0	46.1	8.3	1.5	0.5	--	0.1	10.4
Non-Hispanic black	388	--	0.5	34.4	48.4	13.8	2.1	0.3	0.2	0.4	16.8
Non-Hispanic white	467	--	--	46.5	47.1	5.6	0.5	0.2	--	--	6.3
Seniors											
Mexican-American	863	--	--	0.1	0.6	41.4	47.2	9.2	1.0	0.5	10.7
Cuban	122	--	--	--	0.3	44.4	44.2	9.0	--	2.0	11.0
Puerto Rican	115	--	--	1.9	2.3	30.1	54.9	7.9	0.4	2.5	10.8
Other Latin American	285	--	--	0.7	2.3	41.9	48.3	6.8	0.1	--	6.9
Non-Hispanic black	376	--	--	0.1	0.4	42.8	47.3	7.8	1.2	0.4	9.4
Non-Hispanic white	471	--	--	--	0.5	45.9	49.9	3.4	0.2	0.2	3.8

NOTE:
Percentages are weighted.

1/ Total percent of students whose age is at least two years above the modal age for the grade in the population as a whole (sophomore modal age = 15; seniors = 17). This column is not included in the percent distribution.

Table B.2.—Percent distribution of age by population subgroup for females: Spring 1980

Subgroup	Sample size	13 or younger	14	15	16	17	18	19	20	21 and older	At least 2 years delayed 1/
Sophomores											
Mexican-American	1026	0.3	1.1	42.8	45.3	8.7	1.2	0.2	0.2	0.2	10.5
Cuban	168	--	--	44.9	41.8	13.0	0.2	--	--	--	13.2
Puerto Rican	189	1.2	1.4	46.3	43.4	6.1	1.6	--	--	--	7.7
Other Latin American	342	--	0.4	44.4	46.9	7.5	0.8	--	--	--	8.3
Non-Hispanic black	487	--	0.4	51.1	38.4	7.7	1.8	0.1	--	0.4	10.0
Non-Hispanic white	495	--	0.5	54.2	41.9	3.4	0.0	--	--	--	3.4
Seniors											
Mexican-American	940	--	--	--	1.5	47.2	42.4	7.6	1.0	0.4	9.0
Cuban	208	--	--	--	2.7	52.1	43.0	1.7	0.5	--	2.2
Puerto Rican	177	--	--	--	1.5	49.3	35.1	11.7	2.2	0.2	14.4
Other Latin American	301	--	-	0.2	0.2	48.6	39.8	8.2	2.5	0.5	11.2
Non-Hispanic black	528	--	--	0.7	2.0	51.7	40.4	3.5	1.2	0.6	5.2
Non-Hispanic white	503	--	--	--	1.3	60.3	37.0	1.0	0.3	0.1	1.4

NOTE:
Percentages are weighted.

1/ Total percent of students whose age is at least two years above the modal age for the grade in the population as a whole (sophomore modal age = 15; seniors = 17). This column is not included in the percent distribution.

BEST COPY AVAILABLE

REFERENCES

ASPIRA. 1976. "Social Factors in Educational Attainment Among Puerto Ricans in U.S. Metropolitan Areas, 1970." Puerto Ricans and Education Report #1, September 1976.

Baral, David P. 1979. "Academic Achievement of Recent Immigrants from Mexico." NABE Journal. 3 (Spring):1-13.

Baratz, Jonah. 1973. "Language Abilities of Black Americans." in Kent S. Miller and Ralph Mason Dreger (eds.) Black and Whites in the United States. New York: Seminar Press.

Blau, Peter M. and Duncan, Otis A. 1967. The American Occupational Structure. New York: John Wiley and Sons.

Carter, Thomas P. and Segura, Roberto D. 1979. Mexican Americans in School: A Decade of Change. New York: College Entrance Examination Board.

Duncan, Beverly. 1965. "Dropouts and the Unemployed." Journal of Political Economy. 73(2):121-34.

Erbring, Lutz. 1981. "Cognitive Outcomes in Secondary Education." National Opinion Research Center.

Gomez-Day, Barbara. 1980. "Hispanic Youth and Education" in A Profile of Hispanic Youth. Youth Knowledge Development Report 10.2.

Hernandez, Jose; Estrada, Leo, and Alvirez, David. 1973. "Census Data and the Problem of Conceptually Defining the Mexican American Population," Social Science Quarterly, 53:(4):671-87.

Kimball, William Lloyd. 1968. "Parent and Family Influences on Academic Achievement Among Mexican American Students." Ph.D dissertation. Department of Education, University of California at Los Angeles.

National Center for Education Statistics. 1980. The Condition of Education for Hispanic Americans. Washington D.C.: U.S. Goverment Printing Office.

National Center for Education Statistics, n.d. "Grade Completion of Persons 14 to 25 Years Old. . ." Special Unpublished Table based on SIE, Spring, 1976.

National Opinion Research Center. 1981. "High School and Beyond: Sample Design Report." A report to the National Center for Education Statistics.

Nielsen, Francois. 1980. "Hispanic Youth in U.S. Schools: A Design for Analysis." A report to the National Center for Education Statistics.

Otto, Luther B. and Haller, Archibald O. 1979. "Evidence for a Social Psychological View of the Status Attainment Process." Social Forces. 57(3):887-914.

Portes, Alejandro, McLeod Jr., and Parker, Robert N. 1978. "Immigrant Aspirations." Sociology of Education. 51(October):241-60.

Rogg, Eleanor M. 1974. The Assimilation of Cuban Exiles. New York: Aberdeen Press.

Rosen, Bernard and Aneshensel, Carol. 1978. "Sex Differences in the Educational-Occupational Expectation Process." Social Forces. 57(September).

Smith, Tom W. 1980. "Ethnic Measurement and Identification." Ethnicity. 7:78-95.

U.S. Commission on Civil Rights. 1978. Social Indicators of Equality for Minorities and Women. Report, 1978.

Veltman, Calvin. 1979. "The Assimilation of American Language Minorities: Structure, Pace, and Extent." Report to the National Center for Educational Statistics.

Weinberg, Meyer. 1977. Minority Students: A Research Appraisal. Washington D.C.: Department of Health, Education and Welfare.

GPO 887-422